THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Etude des propriétés de la matrice Jacobienne généralisée dans le cas de fonctions localement Lipschitziennes. Particularisation des résultats à la matrice Hessienne généralisée pour les fonctions de classe C1,1

NICOLAS, Eric

Award date:
1991

Awarding institution:
Universite de Namur

Facultés Universitaires Notre-Dame de la Paix
Namur
Faculté des sciences

Etude des propriétés de la matrice Jacobienne généralisée pour le cas des fonctions localement Lipschitziennes. Particularisation des résultats à la matrice Hessienne généralisée pour les fonctions de classe $C^{1,1}$.

Mémoire présenté pour l'obtention du grade de Licencié en Sciences mathématiques
par

Promoteur:
M. Jean-Jacques STRODIOT

Eric NICOLAS

Année Académique 1990 - 1991

À mes parents.

Merci à Monsieur Jean-Jacques Strodiot
pour l'aide apportée à l'élaboration
de ce mémoire.

CONTENU

INTRODUCTION

La matrice Hessienne d'une fonction f à valeurs réelles et deux fois continûment différentiable joue un rôle important en analyse classique. Par exemple, elle permet de déterminer le développement de Taylor d'ordre un de f, ou de donner des conditions nécessaires pour qu'un point soit un extrémum de f. D'un point de vue algorithmique, elle est utilisée dans la méthode de Newton, ou approximée dans les méthodes quasi-Newton, pour rechercher un extrémum local de f.

Cependant, dans la pratique, les fonctions qui ne sont pas deux fois continûment différentiables se rencontrent fréquemment. Si nous voulons étudier et utiliser de telles fonctions, il est donc nécessaire de trouver un équivalent de la matrice Hessienne.

La classe de fonctions qui sera considérée est celle des fonctions de classe $C^{1,1}$, c'est-à-dire les fonctions continûment différentiables dont le gradient est localement Lipschitzien.

Dans les méthodes d'optimisation, de telles fonctions apparaissent naturellement, même lorsque les fonctions décrivant le problème sont deux fois continûment différentiables. C'est le cas notamment si nous utilisons la méthode du Lagrangien augmenté.

Le gradient d'une fonction de classe $C^{1,1}$ étant localement Lipschitzien, nous commencerons par généraliser la notion de matrice Jacobienne pour une fonction localement Lipschitzienne définie sur $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}^q$. Dans ce but, nous utiliserons l'approche de F.H. CLARKE (références [1] et [2]). Le principal article qui sera étudié est celui de J.-B. HIRIART-URRUTY (référence [6]).

Ensuite, en appliquant les résultats obtenus au gradient d'une fonction de classe $C^{1,1}$, nous généraliserons la notion de matrice Hessienne. Nous en déduirons une régle de différentiation en chaîne qui permettra d'étendre le développement de Taylor d'ordre un. Nous donnerons également des conditions nécessaires d'optimalité du second-ordre pour des problèmes décrits par des fonctions de classe $C^{1,1}$. L'article de J.-B. HIRIART-URRUTY, J.-J. STRODIOT et V.H. NGUYEN (référence [7]) sera à la base de ce développement.

Le premier chapitre est consacré à des rappels d'analyse et de topologie, et le second à un rappel des propriétés du gradient généralisé d'une fonction localement Lipschitzienne définie sur un espace de Banach X et à valeurs réelles.

La définition et les propriétés de la matrice Jacobiennne généralisée sont développées dans le troisième chapitre. Le cas particulier de la matrice Hessienne généralisée fait l'objet du quatrième chapitre.

Pour la clarté de la rédaction, un certain nombre de résultats techniques ont été démontrés dans une annexe.

OUTILS THEORIQUES

Nous allons rappeler ici différentes notions dont nous aurons besoin dans les prochains chapitres. Afin de ne pas alourdir ce travail, les notions les plus connues, notamment celles concernant la convergence, la continuité ou la différentiablité..., ne seront pas énoncées. Et de plus, par la suite, nous ferons appel explicitement aux propriétés de ce chapitre uniquement pour les plus caractéristiques d'entre elles.

1.1 DIVERSES NOTATIONS UTILISÉES.

La plupart du temps, nous travaillerons avec l'espace réel à n dimensions, c'est-à-dire $\mathbb{R}^n$, où n est un nombre naturel fini et supérieur à un. Pour distinguer des espaces réels de dimensions différentes, nous utiliserons les lettres p, q, ...

Lorsque nous voudrons mettre en évidence les composantes d'un point x de $\mathbb{R}^n$, nous noterons ce point x par $(x_i)_{i=1}^{n}$. Ainsi, x_i désignera la $i^{ème}$ composante du point x de $\mathbb{R}^n$.

Un vecteur x de $\mathbb{R}^n$ est un vecteur colonne. Comme nous considérerons également les points de $\mathbb{R}^n$ comme des vecteurs lignes, nous aurons besoin de l'opérateur de transposition que nous noterons par $(\cdot)^T$. Ce même signe de transposition sera utilisé lorsque nous manipulerons des matrices.

Toutes les normes étant équivalentes sur $\mathbb{R}^n$, nous pouvons choisir celle qui nous convient. Choisissons la norme Euclidienne que nous noterons par $\|\cdot\|$. Le produit scalaire associé à cette norme sera noté par $\langle\cdot,\cdot\rangle$.

Occasionnellement, nous considérerons l'espace de Banach X, c'est-à-dire que X désigne un espace normé et complet.

Pour une fonction f définie sur une partie D de $\mathbb{R}^p$ et à valeurs réelles, nous serons amené à considérer, s'il existe, le gradient de f en un point a intérieur à D. Nous le noterons par $\nabla f(a)$. Nous utiliserons également, pour autant quelle existe, la matrice Hessienne de f au point a qui sera notée $\nabla^2 f(a)$.

Pour une fonction F définie sur une partie E de $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}^q$, nous considérerons, si elle existe, la matrice Jacbonienne de F au point a intérieur à E que nous noterons par JF(a).

Les différentes définitions et propriétés que nous allons exposer peuvent être définies pour des espaces plus larges que $\mathbb{R}^n$. Mais, comme nous travaillerons uniquement en dimension finie, les énoncés seront restreints à $\mathbb{R}^n$.

1.2 ELÉMÉNTS TOPOLOGIQUES.

Soit x un point de $\mathbb{R}^n$ et r un nombre réel positif.

Définition 1.2.1: Une partie C de $\mathbb{R}^n$ est <u>compacte</u> si tout recouvrement de C admet un sous-recouvrement fini.

Note 1.2.2: L'espace $\mathbb{R}^n$ étant de dimension finie, la partie C de $\mathbb{R}^n$ est <u>compacte</u> si elle est fermée et bornée.

Théorème 1.2.3 [Théorème de Cantor]

Soit $(\mathcal{F}_n)_{n=1}^{\infty}$ une suite décroissante d'ensembles fermés non vides de $\mathbb{R}^n$.

Si le diamètre de $\mathcal{F}_n$ converge vers 0 pour n tendant vers l'infini,

alors l'intersection de ces ensembles est non vide et est constituée d'un seul élément.

Théorème 1.2.4 [Théorème de Tichonoff]

Tout espace produit d'espaces compacts est compact.

1.3 LIMITES SUPÉRIEURES.

Soit $(x_k)_{k=1}^{\infty}$ une suite de nombres réels et soit f une fonction définie sur une partie D de $\mathbb{R}^p$ et à valeurs réelles.

Définition 1.3.1 : La <u>limite supérieure de la suite</u> $(x_k)_{k=1}^{\infty}$ est le nombre réel défini par $\inf_{N \in \mathbb{N}_0} \sup_{k > N} x_k$.

Nous la noterons par $\limsup_{k \to \infty} x_k$.

Propriété 1.3.2: La limite supérieure de la suite $(x_k)_{k=1}^{\infty}$ est la plus grande valeur d'adhérence de cette suite.

Propriété 1.3.3: Soient $(x_k)_{k=1}^{\infty}$ et $(y_k)_{k=1}^{\infty}$ deux suites de nombres réels,

alors $\limsup_{k \to \infty} (x_k + y_k)$

$$\leq \limsup_{k \to \infty} x_k + \limsup_{k \to \infty} y_k.$$

Définition 1.3.4: La <u>limite supérieure de la fonction</u> f au point a intérieur à D est le nombre réel $\inf_{\varepsilon > 0} \sup_{\|x - a\| < \varepsilon} f(x)$.

Nous la noterons par $\limsup_{x \to a} f(x)$.

Propriété 1.3.5: Soient f et g deux fonctions définies sur une partie D de $\mathbb{R}^p$ et à valeurs réelles et soit a un point intérieur à D,

alors $\limsup\limits_{x\to a} (f(x) + g(x))$

$$\leq \limsup_{x\to a} f(x) + \limsup_{x\to a} g(x).$$

Propriété 1.3.6: Pour qu'une suite réelle soit convergente, ou pour qu'une fonction à valeurs réelles admette une limite en un point, il faut et il suffit que la limite inférieure soit égale à la limite supérieure.

1.4 APPLICATION LOCALEMENT BORNÉE ET SEMI-CONTINUE SUPÉRIEUREMENT.

Soit $\mathcal{M}$ une application multivoque définie sur $\mathbb{R}^p$ et à valeurs dans $\mathcal{P}(\mathbb{R}^q)$, où $\mathcal{P}(\cdot)$ désigne l'ensemble des parties.

Définition 1.4.1 : Soit x un point de $\mathbb{R}^p$. L'application $\mathcal{M}$ est <u>localement</u> <u>bornée</u> s'il existe des constantes réelles $\varepsilon > 0$ et $K > 0$ telles que $\|y\| \leq K$ pour tout z dans $B(x,\varepsilon)$ et pour tout y dans $\mathcal{M}(z)$.

Soit f une fonction définie sur une partie D de $\mathbb{R}^p$ et à valeurs réelles et soit F une fonction définie sur une partie E de $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}^q$.

Cas particuliers 1.4.2: Les applications ∇f et JF sont localement bornées.

Définition 1.4.3: L'application $\mathcal{M}$ est <u>semi-continue</u> <u>supérieurement</u> si pour toute suite $(x_n)_{n=1}^{\infty}$ de points de $\mathbb{R}^p$ qui converge vers un point x de $\mathbb{R}^p$ et pour toute suite $(y_n)_{n=1}^{\infty}$ convergente vers y avec y_n appartenant à $\mathcal{M}(x_n)$, le point y appartient à $\mathcal{M}(x)$.

1.5 Dérivées.

Soit F une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}^q$.

Définition 1.5.1: La fonction F est <u>strictement</u> <u>différentiable</u> au point a de $\mathcal{O}$ s'il existe une matrice de type (q,p), notée $\mathcal{D}F(a)$, telle que, pour tout u dans $\mathbb{R}^p$,

$$\lim_{\substack{x \to a \\ \lambda \to 0^+}} \frac{F(x + \lambda \cdot u) - F(x)}{\lambda} = \mathcal{D}F(a) \cdot u.$$

Propriété 1.5.2: Si F est continûment différentiable au point a de $\mathcal{O}$, alors F est strictement différentiable en a et $JF(a) = \mathcal{D}F(a)$.

Soit f une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles.

Définition 1.5.3: La fonction f admet une <u>dérivée directionnelle</u> au point x de $\mathcal{O}$ dans la direction d de $\mathbb{R}^p$, noté $f'(x;d)$, si la limite suivante existe

$$f'(x;d) = \lim \frac{f(x+\lambda d) - f(x)}{\lambda}.$$

Propriété 1.8.4: Si f est différentiable au point a de $\mathcal{O}$,

alors *i)* pour tout point d de $\mathbb{R}^p$, f est admet une dérivée directionnelle au point a dans la direction d, et

ii) $f'(a;d) = \langle \nabla f(a), d \rangle$.

1.6 Fonctions localement Lipschitziennes.

Soient F et G deux fonctions définies sur un ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}^q$.

Définition 1.6.1 : La fonction F est <u>localement Lipschitzienne</u> au point x_o de $\mathcal{O}$ s'il existe des constantes réelles $\varepsilon > 0$ et $K > 0$ telles que $\| F(x) - F(y) \| \leq K \cdot \| x - y \|$ pour tout x et y dans la boule $B(x_o, \varepsilon)$.

Propriété 1.6.2: Soit x_o un point de $\mathcal{O}$. Si F et G sont des fonctions localement Lipschitziennes en x_o, alors $(F + G)$ est une fonction localement Lipschitzienne au point x_o.

Propriété 1.6.3: Soit x_o un point de $\mathcal{O}$ et soit r un nombre réel. Si F est une fonction localement Lipschitzienne en x_o, alors $(r \cdot F)$ est une fonction localement Lipschitzienne en x_o.

Propriété 1.6.4: Soit x_o un point de $\mathcal{O}$. Si F est une fonction localement Lipschitzienne en x_o, alors F est continue en x_o.

Propriété 1.6.5 [Théorème de Rademacher]

Si F est localement Lipschitzienne en tout point de $\mathcal{O}$, alors l'ensemble Ω_F des points x de $\mathcal{O}$ où F est non différentiable est de mesure de Lebesgue nulle.

Remarque: Une démonstration de ce théorème se trouve dans le chapitre 4 de [4].

1.7 ELÉMENTS D'ANALYSE CONVEXE.

Soit S une partie de $\mathbb{R}^p$.

Définition 1.7.1 : La partie S de $\mathbb{R}^p$ est convexe si pour tout x et y dans S et pour tout nombre λ de l'intervalle réel $[0,1]$, le point $(\lambda \cdot x + (1-\lambda) \cdot y)$ appartient à S.

Propriété 1.7.2 [Théorème de séparation stricte des convexes]

Soient A et B deux parties convexes, non vides et disjointes de $\mathbb{R}^p$.

Si A est un fermé et si B est un compact,

alors il existe un nombre réel $\alpha > 0$ et un point u de $\mathbb{R}^p$ tels que

$$\langle x,u\rangle < \alpha < \langle y,u\rangle$$

pour tout x dans A et Y dans B.

Définition 1.7.3: La <u>fonction</u> <u>de</u> <u>support</u> de la partie S non vide, convexe et fermée est la fonction s définie sur $\mathbb{R}^p$ par

$$s(u) = \sup \{ \langle x,u\rangle \mid x \in S \} \text{ pour tout } u \text{ dans } \mathbb{R}^p.$$

Propriété 1.7.4: Soient A et B deux parties non vides, convexes et fermées de $\mathbb{R}^p$.

Si $\sup \{ \langle x,u\rangle \mid x \in A \} \leq \sup \{ \langle x,u\rangle \mid x \in A \}$

pour tout u appartenant à $\mathbb{R}^p$,

alors A est inclus dans B.

Définition 1.7.5: L'<u>enveloppe</u> <u>convexe</u> de la partie S de $\mathbb{R}^p$, noté co(S), est le plus petit ensemble convexe contenant S.

Propriété 1.7.6 [Caractérisation de l'enveloppe convexe]

Les assertions suivantes sont équivalentes.

i) x appartient à co(S),

ii) il existe un nombre k de $\mathbb{N}_0$ tel que

$$\sum_{i=1}^{k} \lambda_i \cdot x_i,$$

où x_i appartient à S et $\lambda_i \geq 0$ pour tout i compris entre 1 et k, et $\sum_{i=1}^{k} \lambda_i = 1$.

Propriété 1.7.7: Si S est une partie compacte de $\mathbb{R}^P$, alors co(S) est une partie compacte de $\mathbb{R}^P$.

Propriété 1.7.8: Pour tout u appartenant à $\mathbb{R}^P$,

$$\sup \{ \langle x,u \rangle \mid x \in S \} = \sup \{ \langle x,u \rangle \mid x \in co(S) \}.$$

Soit f une fonction définie sur la partie S de $\mathbb{R}^P$ et à valeurs réelles.

Définition 1.7.9: La <u>fonction</u> f est <u>convexe</u> sur S si

i) S est une partie convexe de $\mathbb{R}^P$, et si

ii) $f(\lambda \cdot x + (1-\lambda) \cdot y) \leq \lambda \cdot f(x) + (1-\lambda) \cdot f(y)$

pour tout x et y appartenant à S et pour tout nombre λ de l'intervalle réel $]0,1[$.

Propriété 1.7.10: Si S est une partie ouverte et convexe de $\mathbb{R}^p$ et si f est différentiable sur S, alors les assertions suivantes sont équivalentes:

i) f est une fonction convexe sur S,

ii) $\langle \nabla f(y) - \nabla f(x) \, , \, y - x \rangle \geq 0$

pour tout x et y appartenant à S.

Propriété 1.7.11: Si S est une partie ouverte et convexe de $\mathbb{R}^p$ et si f est deux fois continûment différentiable sur S, alors les assertions suivantes sont équivalentes:

i) f est une fonction convexe sur S,

ii) la matrice Hessienne de f est semi-définie positive en tout point de S.

Définition 1.7.12: Le <u>cône tangent de Bouligant</u> à la partie S au point a de S est l'ensemble, noté $T_S(a)$, des points d de $\mathbb{R}^p$ pour lesquels il existe une suite $(d_k)_{k=1}^{\infty}$ dans $\mathbb{R}^p$ qui converge vers d, et il existe une suite $(\alpha_k)_{k=1}^{\infty}$ de nombres réels qui converge vers 0 par valeurs positives telles que $(a + \alpha_k \cdot d_k)$ appartient à S quel que soit $k \geq 1$.

Autrement dit,

$$T_S(a) = \Big\{ d \in \mathbb{R}^p \,\Big|\, \exists (d_k)_{k=1}^{\infty} \subseteq \mathbb{R}^p \text{ convergente vers } d \text{ et } \exists (\alpha_k)_{k=1}^{\infty} \subseteq \mathbb{R} \text{ convergente vers 0 par valeurs positives telles que } (a + \alpha_k \cdot d_k) \in S \ \forall k \geq 1 \Big\}.$$

1.8 CONDITIONS D'OPTIMALITÉ.

Soit f une fonction défine sur un ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles.

Propriété 1.8.1 [Condition nécessaire du premier ordre]

Supposons que f admette toutes les dérivées partielles du premier ordre au point a de $\mathcal{O}$.
Si a est un point d'extrémum local de f,
alors le gradient de f au point a s'annule.

Propriété 1.8.2 [Condition nécessaire du second ordre]

Supposons que f soit deux fois continûment différentiable sur une boule ouverte centrée en un point a de $\mathcal{O}$.
Si a est un point de minimum local de f,
alors *i)* le gradient de f s'annule au point a,
ii) la matrice Hessienne de f au point a est semi-définie positive.

Soient g_i et h_j des fonctions définies sur $\mathbb{R}^p$ et à valeurs réelles quels que soient i dans un ensemble fini de nombres naturels I et j dans un ensemble fini de nombres naturels E. Notons par m, respectivement par n, le nombre d'éléments de I, respectivement de E.

Considérons la forme générale d'un problème avec contraintes,

$$(PAC)\begin{cases}\text{Minimiser } f(x) \\ \text{sous conditions: } g_i(x) \leq 0 \text{ pour tout } i \text{ dans } I \\ \qquad\qquad\qquad\quad h_j(x) = 0 \text{ pour tout } j \text{ dans } E \\ \qquad\qquad\qquad\quad x \text{ appartient à } \mathcal{O}.\end{cases}$$

Pour ce problème, nous avons les conditions nécéssaires suivantes.

Propriété 1.8.3 [Conditions nécessaires d'optimalité de Fritz-John]

Soit a un point de minimim local du problème (PAC).

Si les fonctions f, g_i et h_j sont continûment différentiables sur leur domaine respectif, alors, il existe un vecteur non nul de $\mathbb{R}^{m+n+1}$, noté $(\lambda_o,\lambda_1,\ldots,\lambda_m,\mu_1,\ldots,\mu_n)$, tel que

$$(FJ)\begin{cases}\lambda_o\cdot\nabla f(x_o) + \sum\limits_{i\in I}\lambda_i\cdot\nabla g_i(x_o) + \sum\limits_{j\in E}\mu_j\cdot\nabla h_j(x_o) = 0, \\ \lambda_o \geq 0, \text{ et pour tout } i \text{ dans } I, \\ \lambda_i \geq 0 \text{ et } \lambda_i\cdot g_i(x_o) = 0.\end{cases}$$

En ajoutant une condition de régulatité du premier ordre, appelée <u>condition de qualification des contraintes</u>, les conditions nécessaires d'optimalité de Fritz-John conduisent à la propriété suivante.

Propriété 1.8.4 [Conditions nécessaires d'optimalité de Kuhn-Tucker]

Soit a un point de minimim local du problème *(PAC)*.

Si les fonctions f, g_i et h_j sont continûment différentiables sur leur domaine respectif, et si la condition de qualification des contraintes est vérifiée au point a,

alors, il existe un vecteur de $\mathbb{R}^{m+n}$, noté $(\lambda_1,\ldots,\lambda_m,\mu_1,\ldots,\mu_n)$, tel que

$$(KT)\quad \begin{cases} \nabla f(x_0) + \sum\limits_{i\in I} \lambda_i \nabla g_i(x_0) + \sum\limits_{j\in E} \mu_j \cdot \nabla h_j(x_0) = 0, \\ \lambda_0 \geq 0, \text{ et pour tout } i \text{ dans } I, \\ \lambda_i \geq 0 \text{ et } \lambda_i \cdot g_i(x_0) = 0. \end{cases}$$

Exemple 1.8.5 de condition de qualification des contraintes:

Soit I(a), l'ensemble des points de I pour lesquels les contraintes d'inégalité sont actives au point a.

Si les vecteurs $\nabla g_i(a)$ et $\nabla h_j(a)$ sont linéairement indépendants quels que soient i dans I(a) et j dans E,

alors la condition de qualification des contraintes est vérifiée au point a.

1.9 PRINCIPE DE LA BORNE UNIFORME.

Définition 1.9.1 : La <u>norme</u> d'un <u>opérateur</u> T, notée par $\|T\|$, dont le domaine D est dans l'espace X et à valeurs dans Y est définie par $\|T\| = \sup_{\substack{x \in D \\ \|x\| \le 1}} \|T(x)\|$.

Propriété 1.9.1 [Principe de la borne uniforme]

Soient X un espace de Banach et Y un espace normé. Et soit $(T_i)_{i \in I}$ une famille d'opérateurs linéaires bornés définis sur X et à valeurs dans Y.

Si $\sup_{i \in I} \|T_i(x)\|$ est fini quel que soit x dans X, alors $\sup_{i \in I} \|T_i\|$ est fini.

1.10 PROPRIÉTÉS DE L'INTÉGRALE.

Soient f et g deux fonctions définies sur un intervalle réel [a,b] et à valeurs réelles. Supposons qu'elles soient intégrables sur l'intervalle [a,b].

Propriété 1.10.1 : Si $f(x) \le g(x)$ pour tout x appartenant à [a,b],

$$\text{alors } \int_a^b f(x)\,dx \le \int_a^b g(x)\,dx.$$

Soient u et v deux fonctions définies sur l'intervalle réel $[\alpha,\beta]$ et à valeurs dans l'intervalle réel $[a,b]$ telles que $u(x) \leq v(x)$ pour tout x dans $[\alpha,\beta]$. Et soit Φ une fonction définie sur $[\alpha,\beta] \times [a,b]$ et à valeurs réelles.

Supposons que Φ soit intégrable sur son domaine, et posons

$$I(x) = \int_{u(x)}^{v(x)} \Phi(t,x)\, dt.$$

Propriété 1.10.2: Si les fonctions u,v et f sont continues sur leur domaine respectif,
alors I est une fonction continue sur $[a,b]$.

Propriété 1.10.3: Si *i)* Φ est une fonction continue sur son domaine,
ii) Φ est dérivable par rapport à x et que la dérivée Φ'_x est continue,
iii) u et v sont dérivables,
alors I est dérivable et

$$I'(x) = \int_{u(x)}^{v(x)} \Phi'_x(t,x)\, dt + \Phi(u(x),x)\cdot u'(x) + \Phi(v(x),x)\cdot v'(x).$$

GRADIENT GENERALISE

Nous allons rappeler quelques propriétés du gradient généralisé que nous utiliserons par la suite. Pour cette raison, seules les caractéristiques intéressantes pour les chapitres suivants seront énoncées. Les propriétés décrites dans ce chapitre étant des outils pour la suite, nous n'en développerons pas les démonstrations. Les preuves se trouvent dans le livre [2] de F.H. CLARKE.

Considérons X un espace de Banach et X' l'espace dual, c'est-à-dire l'ensemble des fonctionnelles linéaires continues définies sur X. Et soit f une fonction définie sur un ouvert $\mathcal{O}$ de X et à valeurs réelles. Nous supposerons que f est localement Lipschitzienne en tout point de $\mathcal{O}$.

2.1 DÉRIVÉE DIRECTIONNELLE GÉNÉRALISÉE.

La fonction f étant localement Lipschitzienne, l'existence d'une dérivée directionnelle de f n'est plus assurée.

Cette constatation motive la définition suivante.

Définition 2.1.1 : La <u>dérivée</u> <u>directionnelle</u> <u>généralisée</u> de f au point x_o de $\mathcal{O}$ dans la direction u de X est la fonction, notée $f^{\circ}(\cdot;\cdot)$, définie par

$$f^{\circ}(x_o;u) = \limsup_{\substack{y \to x \\ \lambda \to o^+}} \frac{f(y+\lambda u) - f(y)}{\lambda}.$$

2.2 GRADIENT GÉNÉRALISÉ.

Nous allons commencer par présenter des propriétés vraies sur tout espace de Banach. Ensuite, nous verrons des propriétés particulières dans le cas où l'espace envisagé est $\mathbb{R}^p$.

Définition 2.2.1 : Le <u>gradient</u> <u>généralisé</u> de f au point x_o de $\mathcal{O}$ est le sous-ensemble de X', noté $\partial f(x_o)$, défini par

$$\partial f(x_o) = \{ x \in X' \mid f^{\circ}(x_o;u) \geq \langle x,u\rangle \text{ pour tout } u \in X \},$$

où $\langle\cdot,\cdot\rangle$ est l'application linéaire définie sur $X' \times X$ et à valeurs réelles par $\langle x,u\rangle = x(u)$ pour tout $x \in X'$ et $u \in X$.

Pour cette définition, nous énonçons les propriétés qui suivent.

Propriété 2.2.2: Pour tout x_o appartenant à $\mathcal{O}$,
$\partial f(x_o)$ est un ensemble non vide, convexe et faiblement compact.

Propriété 2.2.3: Soit x_o appartenant à $\mathcal{O}$.
Si f est continûment différentiable en x_o,
alors $\partial f(x_o) = \{ \nabla f(x_o) \}$.

Propriété 2.2.4: Pour tout x_o appartenant à $\mathcal{O}$ et u dans X,

$$f^{\circ}(x_o;u) = \max \{ \langle x,u \rangle \mid x \in \partial f(x_o) \}.$$

Autrement dit $f^{\circ}(x_o;\cdot)$ est la fonction de support de $\partial f(x_o)$.

Propriété 2.2.5: Soit x_o appartenant à $\mathcal{O}$,
alors $\partial(s\cdot f)(x_o) = s\cdot\partial f(x_o)$ pour tout s dans $\mathbb{R}$.

Propriété 2.2.6: Soit x_o appartenant à $\mathcal{O}$ et soit $(f_i)_{i=1}^{n}$ une famille finie de fonctions localement Lipschitziennes définie sur $\mathcal{O}$ et à valeurs réelles,
alors $\partial\left(\sum_{i=1}^{n} f_i\right) \subseteq \sum_{i=1}^{n} \partial f_i(x_o)$.

Propriété 2.2.7: Soit x_o appartenant à $\mathcal{O}$,
si x_o est un point d'extrémum local pour la fonction f, alors 0 appartient à $\partial f(x_o)$.

Soient x et y appartenant à X, posons

$$[x,y] = \{\ z \in X \mid \exists\ \lambda \text{ dans l'intervalle réel } [0,1] \text{ tel que } z = \lambda\cdot x + (1-\lambda)\cdot y\ \},$$

et,

$$(x,y) = \{\ z \in X \mid \exists\ \lambda \text{ dans l'intervalle réel }]0,1[\text{ tel que } z = \lambda\cdot x + (1-\lambda)\cdot y\ \}.$$

Théorème 2.2.8 [Théorème de la valeur intermédiaire]

Soient x et y deux points de X, supposons que [x,y] soit contenu dans $\mathcal{O}$, alors il existe un point u de (x,y) tel que

$f(y) - f(x)$ appartient à $\langle \partial f(u), y - x\rangle$.

Dans les chapitres qui suivent, nous travaillerons en dimension finie. Ces propriétés restent évidemment correctes mais nous aurons $X = \mathbb{R}^p$ et $X' = \mathbb{R}^p$.

Ainsi, les propriétés faibles deviennent fortes, notamment $\partial f(x_o)$ est compact. De plus, $\langle\cdot,\cdot\rangle$ est alors le produit scalaire sur $\mathbb{R}^p$.

Propriété 2.2.9: L'application

$$\begin{aligned} \partial f : \mathbb{R}^p &\longrightarrow \mathcal{P}(\mathbb{R}^p) \\ x &\longrightarrow \partial f(x) \end{aligned}$$

est *i)* localement bornée

ii) semi-continue supérieurement.

Les démonstrations de ces propriétés ne se trouvent pas dans la référence [2] de F.H. CLARKE. Néanmoins, nous n'envisagerons pas la preuve de ces propriétés car, dans le chapitre suivant, nous développerons des propriétés similaires dont les preuves se déroulent selon le même schéma.

A présent, nous allons voir des caractérisations du gradient généralisé et de la dérivée directionnelle généralisée qui sont applicables uniquement sur l'espace de dimension finie $\mathbb{R}^p$.

Notons par Ω_f l'ensemble des points de $\mathcal{O}$ où f n'est pas différentiable.

Théorème 2.2.10: Soit x_o dans $\mathcal{O}$ et soit S inclus dans $\mathbb{R}^p$ de mesure de Lebesgue nulle, alors

$$\partial f(x_o) = \mathrm{co}\ [\partial(f,x_o)],$$

où $\partial(f,x_o) = \{\ x \in \mathbb{R}^p \mid \exists (x_n)_{n=1}^{\infty} \subseteq S^c \cap \Omega_f^c$ convergente vers x_o telle que $x = \lim\limits_{n\to\infty} \nabla f(x_n)\ \}$

De ce théorème, nous déduisons la caractérisation de la dérivée directionnelle généralisée.

Corollaire 2.2.11: Soient x_o appartenant à $\mathcal{O}$ et u dans $\mathbb{R}^p$,

$$\text{alors } f^{\circ}(x_o;u) = \limsup_{\substack{y \to x \\ y \in S^c \cap \Omega_f^c}} \langle \nabla f(y), u \rangle.$$

Voici résumées les propriétés concernant le gradient généralisé dont nous aurons besoin pour les prochains chapitres.

MATRICE JACOBIENNE GENERALISEE

Afin de pouvoir définir une extension de la matrice Jacobienne dans le cas d'une fonction localement Lipschitzienne, nous allons développer dans ce chapitre un nouvel objet, la matrice Jacobienne généralisée.

Dans ce but, soit F une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}^q$. Les q composantes de F seront notées par $f_1,\ldots,f_q$. Chaque f_i sera donc une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles. De plus, nous supposerons que F, et donc chaque f_i, est localement Lipschitzienne en tout point de $\mathcal{O}$.

3.1 DÉFINITION ET PREMIÈRES PROPRIÉTÉS.

Notons par $\mathcal{M}(q,p)$ l'ensemble des matrices réelles ayant q lignes et p colonnes et par Ω_F^c le complémentaire dans $\mathcal{O}$ de l'ensemble Ω_F des points où F n'est pas différentiable.

En dimension finie, toutes les normes étant équivalentes, nous pouvons choisir sur $\mathcal{M}(q,p)$ une norme qui nous convient. Sauf indications contraires, pour une matrice M de $\mathcal{M}(q,p)$, nous utiliserons la norme définie de la façon suivante:

$\|M\|=\max\ \{\,|m_{ij}|\ \mid\ 1\le i\le q \text{ et } 1\le j\le p\ \}$, où m_{ij} l'élément situé à la $i^{\text{ème}}$ ligne et à la $j^{\text{ème}}$ colonne de la matrice M.

A présent définissons la matrice Jacobienne généralisée.

Définition 3.1.1 : La <u>matrice</u> <u>Jacobienne</u> <u>généralisée</u> de F au point x_o appartenant à $\mathcal{O}$, notée par $\mathfrak{J}F(x_o)$, est l'enveloppe convexe des matrices M de $\mathcal{M}(q,p)$ telles que $M = \lim\limits_{n\to\infty} JF(x_n)$, où $(x_n)_{n=1}^{\infty}$ est une suite de points de Ω_F^c qui converge vers x_o. Autrement dit,

$$\mathfrak{J}F(x_o)=\mathrm{co}\ [\mathfrak{J}(F,x_o)]$$

$$\text{où } \mathfrak{J}(F,x_o)=\left\{\, M\in\mathcal{M}(q,p)\ \mid\ \exists (x_n)_{n=1}^{\infty}\subseteq \Omega_F^c \text{ convergente vers } x_o \text{ telle que } M=\lim_{n\to\infty} JF(x_n) \right\}. \qquad (3.1)$$

Compte tenu de cette définition, le gradient généralisé est, lorsque q = 1, un cas particulier de la matrice Jacobienne généralisée (à un signe de transposition près). En effet, grâce au théorème 2.2.10, nous obtenons alors que

$$\mathfrak{J}F(x_o) = [\partial F(x_o)]^T.$$

Ainsi, toutes les propriétés que nous rencontrerons pour la matrice Jacobienne généralisée seront applicables au gradient généralisé.

Pour vérification, dans toutes les caractéristiques de la matrice Jacobienne généralisée que nous verrons, nous tâcherons de retrouver ce que nous connaissons à propos du gradient généralisé. Mais, rien ne permet d'affirmer que la réciproque est vraie. Et plus précisément, nous verrons que ces deux objets ont des caractéristiques différentes.

Nous allons décrire quelques propriétés qui seront utilisées par la suite.

Propriété 3.1.2 : Pour tout x_o appartenant à $\mathcal{O}$,

$\mathfrak{J}F(x_o)$ est non vide, compact et convexe.

Preuve : Cette démonstration comporte 3 étapes.

i) $\mathfrak{J}F(x_o)$ est convexe par définition.

ii) Montrons que $\mathfrak{J}F(x_o)$ est non vide.

Nous commençons par remarquer que:

Dans toute boule de centre x_o contenue dans $\mathcal{O}$, il existe un point x où F est différentiable. (3. 2)

En effet, si ce n'est pas le cas, il existe une boule $B(x_o,\varepsilon)$ contenue dans $\mathcal{O}$ telle que, en tout point, F n'est pas différentiable. Ainsi, $B(x_o,\varepsilon)$ est inclus dans Ω_F, et comme $B(x_o,\varepsilon)$ n'est pas de mesure de Lebesgue nulle, alors Ω_F n'est pas de mesure de Lebesgue nulle, ce qui contredit le théorème 1.6.5 de Rademacker.

Maintenant, considérons la suite $(\varepsilon_i)_{i=1}^{\infty} = (\frac{\varepsilon}{i})_{i=1}^{\infty}$. Par application de l'assertion (3.2), il existe une suite $(x_i)_{i=1}^{\infty}$ telle que, pour tout i dans $\mathbb{N}_o$, x_i appartient à $B(x_o,\varepsilon_i)$ et F est différentiable en x_i. Comme, $B(x_o,\varepsilon_j)$ est inclus dans $B(x_o,\varepsilon_i)$ pour tout $i < j$, et que la suite $(\varepsilon_i)_{i=1}^{\infty}$ converge vers 0, nous avons, par application du théorème 1.2.3 de Cantor, que la suite $(x_i)_{i=1}^{\infty}$ converge vers x_o.

De plus, comme l'application JF est localement bornée, la suite $(JF(x_i))_{i=1}^{\infty}$ est bornée. Ainsi, par application du théorème de Bolzano-Weierstrass, il existe une sous-suite $(JF(x_{q_i}))_{i=1}^{\infty}$ extraite de $(JF(x_i))_{i=1}^{\infty}$, qui converge vers une matrice M. Cette matrice vérifie les conditions de définition de $\mathfrak{J}F(x_o)$. Donc, la matrice Jacobienne généralisée est non vide.

iii) Montrons que $\mathfrak{J}F(x_o)$ est compact.

L'espace $\mathcal{M}(q,p)$ étant de dimension finie, il suffit de montrer que $\mathfrak{J}F(x_o)$ est borné et fermé.

a) Montrons que $\mathfrak{J}F(x_o)$ est borné.

Par la propriété 1.7.7, nous savons que l'enveloppe convexe d'un ensemble borné est bornée. Ainsi, il suffit de montrer que $\mathfrak{J}(F,x_o)$ est borné.

Soit M une matrice de $\mathfrak{J}(F,x_o)$.

Par définition de cet ensemble, il existe une suite $(x_n)_{n=1}^{\infty}$, de points de Ω_F^c convergente vers x_o, telle que $M = \lim_{n\to\infty} JF(x_n)$.

Ainsi, nous avons que $\|M\| = \lim_{n\to\infty}\|JF(x_n)\|$.

Comme l'application JF est localement bornée, il existe $K \geq 0$ tel que pour tout x dans Ω_F^c, nous avons $\|JF(x)\| \leq K$.

Il suit que, pour tout nombre n de $\mathbb{N}_0$, $\|JF(x_n)\| \leq K$. Et par passage à la limite, nous obtenons que $\|M\| \leq K$.

b) Montrons que $\mathfrak{J}F(x_0)$ est fermé.

Par la propriété 1.7.7, nous savons que l'enveloppe convexe d'un ensemble fermé est fermée. Il suffit donc de montrer que $\mathfrak{J}(F,x_0)$ est fermé, c'est-à-dire que si M est la limite d'une suite $(M_n)_{n=1}^{\infty}$ de $\mathfrak{J}(F,x_0)$, alors M appartient à $\mathfrak{J}(F,x_0)$.

Pour tout $n \geq 1$, comme M_n appartient à $\mathfrak{J}(F,x_0)$, il existe, par définition, une suite $(x_i^n)_{i=1}^{\infty}$, de points de Ω_F^c convergente vers x_0, telle que $M_n = \lim_{i\to\infty} JF(x_i^n)$.

Construisons la suite $(x_{i_n}^n)_{n=1}^{\infty}$ de la façon suivante:

Soit n un nombre de $\mathbb{N}_0$. Puisque $(x_i^n)_{i=1}^{\infty}$ converge vers x_0 et $(JF(x_i^n))_{i=1}^{\infty}$ converge vers M_n, il existe un nombre i_n de $\mathbb{N}_0$ tel que $\|x_{i_n}^n - x_0\| < \frac{1}{n}$ et $\|JF(x_{i_n}^n) - M_n\| < \frac{1}{n}$. Ainsi, la suite $(x_{i_n})_{n=1}^{\infty}$ converge vers x_0 et, en utilisant l'inégalité triangulaire, la suite $(JF(x_{i_n}^n))_{n=1}^{\infty}$ converge vers la même limite que la suite $(M_n)_{n=1}^{\infty}$, c'est-à-dire vers M.

□

Propriété 3.1.3 : L'application
$$\mathfrak{J}F : \mathcal{O} \subseteq \mathbb{R}^p \longrightarrow \mathcal{P}(\mathcal{M}(q,p))$$
$$x \longrightarrow \mathfrak{J}F(x)$$

est localement bornée.

Preuve :

Soit x_o un point de $\mathbb{R}^p$. Nous devons montrer qu'il existe $\varepsilon > 0$ et $K > 0$ tels que $\|J\| \leq K$ pour tout x appartenant à $B(x_o,\varepsilon)\cap\Omega_F^c$ et pour tout J appartenant à $\mathfrak{J}F(x)$.

L'enveloppe convexe conservant le caractère borné d'un ensemble, il suffit de montrer qu'il existe $\varepsilon > 0$ et $K > 0$ tels que $\|M\| \leq K$ pour tout x dans $B(x_o,\varepsilon)\cap\Omega_F{}^c$ et pour toute matrice M de $\mathfrak{J}(F,x)$.

L'application JF étant localement bornée, il existe $\varepsilon > 0$ et $K > 0$ tels que $\|JF(x)\| \leq K$ pour tout x appartenant à $B(x_o,\varepsilon)\cap\Omega_F^c$.

Soient x dans $B(x_o,\varepsilon)\cap\Omega_F^c$ et M une matrice de $\mathfrak{J}(F,x)$. Par définition de $\mathfrak{J}(F,x)$, il existe une suite $(x_n)_{n=1}^{\infty}$ convergente vers x où, x_n appartient à Ω_F^c pour tout $n \geq 1$, et $M = \lim_{n\to\infty} JF(x_n)$. Comme la suite $(x_n)_{n=1}^{\infty}$ converge vers x, il existe un nombre N de $\mathbb{N}_o$ tel que x_n appartient à $B(x_o,\varepsilon)$ pour tout $n > N$. Ainsi, nous déduisons que x_n appartient à $B(x_o,\varepsilon)\cap\Omega_F^c$. Grâce au caractère localement borné de JF, il suit que, pour tout $n > N$, nous avons $\|JF(x_n)\| \leq K$.

Enfin, par passage à la limite, nous obtenons que

$$\|M\| = \lim_{n\to\infty} \|JF(x_n)\| \leq K.$$

Donc, en prenant ε et K, la propriété est démontrée.

□

Propriété 3.1.4 : L'application $\mathfrak{J}F : \mathcal{O} \subseteq \mathbb{R}^p \longrightarrow \mathcal{P}(\mathcal{M}(q,p))$

$$x \longrightarrow \mathfrak{J}F(x)$$

est semi-continue supérieurement.

Preuve :

Soit x_o un point de $\mathbb{R}^p$. Montrons que, pour toutes suites $(x_n)_{n=1}^{\infty}$ contenue dans $\mathcal{O}$ et convergente vers x_o, et $(M_n)_{n=1}^{\infty}$ contenue dans $\mathcal{M}(q,p)$ et convergente vers M_o telles que M_n appartient à $\mathfrak{J}F(x_n)$ pour tout nombre n de $\mathbb{N}_o$, nous avons que M_o appartient à $\mathfrak{J}F(x_o)$. La démonstration étant identique à celle montrant que $\mathfrak{J}F(x_o)$ est un ensemble fermé, nous ne la détaillerons pas .

□

Si nous envisageons le cas différentiable, retrouvons nous la matrice Jacobienne dans cette définition? La réponse est founie par la propriété suivante.

Propriété 3.1.5 : Si F est strictement différentiable en x_o, alors $\mathfrak{J}F(x_o)=\{ \mathcal{D}F(x_o) \}$, où $\mathcal{D}F(x_o)$ désigne la différentielle stricte de F en x_o.

Comme cette propriété ne sera plus utilisée par la suite, nous ne la démontrerons pas (pour plus de détails, voir [1]).

Dans le cas continûment différentiable, cette propriété nous indique que $\mathfrak{J}F(x_o)=\{ JF(x_o) \}$. En effet, comme F est continûment différentiable en x_o, nous avons, grâce à la propriété 1.5.2, que F est strictement différentiable en x_o et $\mathcal{D}F(x_o) = JF(x_o)$.

3.2 BIFONCTION DE SUPPORT.

Soit x_o un point de $\mathcal{O}$. Afin de caractériser plus facilement $\mathfrak{J}F(x_o)$, nous allons essayer d'exprimer sa fonction de support.

Dans ce but, considérons, dans $\mathcal{M}(q,p)$, le produit scalaire, noté par $\langle\langle \cdot , \cdot \rangle\rangle$ et défini de la manière suivante:

$$\langle\langle A,B \rangle\rangle = \text{trace } (A \cdot B^T), \text{ pour tout } A \text{ et } B \text{ dans } \mathcal{M}(q,p).$$

Autrement dit, si $A = (a_{ij})$ et $B = (b_{ij})$, où $1 \leq i \leq q$ et $1 \leq j \leq p$, alors $\langle\langle A,B \rangle\rangle = \sum_{i=1}^{q} \sum_{j=1}^{p} a_{ij}\ b_{ij}$.

Pour tout U appartenant à $\mathcal{M}(q,p)$, cherchons une forme plus explicite de la fonction de support de $\mathfrak{J}F(x_o)$, c'est-à-dire calculons

$$\max \{ \langle\langle M,U \rangle\rangle \mid M \in \mathfrak{J}F(x_o) \}.$$

Remarque : Ce maximum existe car $\mathfrak{J}F(x_o)$ est un ensemble compact (cfr. propriété 3.1.2) et $\langle\langle \cdot ,U \rangle\rangle$ est une application continue quel que soit la matrice U de $\mathcal{M}(q,p)$.

Commençons par montrer la proposition suivante:

Proposition 3.2.1 : Pour tout U appartenant à $\mathcal{M}(q,p)$,

$$\max \{ \langle\langle M,U \rangle\rangle \mid M \in \mathfrak{J}F(x_o) \} = \limsup_{\substack{x \to x_o \\ x \in \Omega_F^c}} \langle\langle JF(x),U \rangle\rangle.$$

Preuve : Nous allons utiliser les résultats suivants.

Soient ϕ une fonction définie sur $\mathbb{R}^n$ et à valeurs réelles et x appartenant à $\mathbb{R}^n$.

(R1) Pour toute suite $(x_n)_{n=1}^{\infty}$ de points de $\mathbb{R}^n$ qui converge vers x, nous avons que

$$\limsup_{n\to\infty} \phi(x_n) \le \limsup_{y\to x} \phi(y).$$

(La démonstration de ce résultat se trouve à l'annexe B.1.)

(R2) Soit c un nombre réel, alors les deux assertions suivantes sont équivalentes.

(R2.1) $\limsup_{y\to x} \phi(y) \le c$

(R2.2) Pour toute suite $(x_n)_{n=1}^{\infty}$ de points de $\mathbb{R}^n$ qui converge vers x, nous avons que $\limsup_{n\to\infty} \phi(x_n) \le c$.

(La démonstration de ce résultat se trouve à l'annexe B.2.)

Soit U un matrice de $\mathcal{M}(q,p)$.

1°) Montrons que:

$$\max \{ \langle\langle M,U\rangle\rangle \mid M \in \mathfrak{J}F(x_o) \} \le \limsup_{\substack{x \to x_o \\ x \in \Omega_F^c}} \langle\langle JF(x),U\rangle\rangle. \tag{3.3}$$

Soit M appartenant à $\mathfrak{J}F(x_o)$. Par définition de $\mathfrak{J}F(x_o)$, il existe un nombre k de $\mathbb{N}_o$ tel que $M = \sum_{i=1}^{k} \lambda_i M_i$, où, pour tout $1 \le i \le k$, M_i appartient à $\mathfrak{J}(F,x_o)$, $\lambda_i \ge 0$ et $\sum_{i=1}^{k} \lambda_i = 1$.

Soit i tel que $1 \le i \le k$, comme M_i appartient à $\mathfrak{J}(F,x_o)$, il existe une suite $(x_i^n)_{n=1}^{\infty}$ de points de Ω_F^c qui converge vers x_o telle que $M_i = \lim_{n\to\infty} JF(x_i^n)$. Comme l'application $\langle\langle \cdot ,U\rangle\rangle$ est continue, il suit que $\langle\langle M_i,U\rangle\rangle = \lim_{n\to\infty} \langle\langle JF(x_i^n),U\rangle\rangle$.

Ainsi, par application du résultat *(R1)*, nous obtenons

$$\langle\langle M_i, U\rangle\rangle \leq \limsup_{\substack{x \to x_o \\ x \in \Omega_F^c}} \langle\langle JF(x), U\rangle\rangle.$$

Finalement, sachant que $M = \sum_{i=1}^{k} \lambda_i M_i$ et que $\sum_{i=1}^{k} \lambda_i = 1$, nous avons

$$\langle\langle M, U\rangle\rangle \leq \limsup_{\substack{x \to x_o \\ x \in \Omega_F^c}} \langle\langle JF(x), U\rangle\rangle.$$

Ceci étant vérifié quel que soit la matrice M de $\mathfrak{J}F(x_o)$, c'est également correct si nous prenons le maximum sur l'ensemble $\mathfrak{J}F(x_o)$. Donc l'inégalité (3.3) est démontrée.

2°) Montrons que:

$$\limsup_{\substack{x \to x_o \\ x \in \Omega_F^c}} \langle\langle JF(x), U\rangle\rangle \leq \max \{ \langle\langle M, U\rangle\rangle \mid M \in \mathfrak{J}F(x_o) \}. \qquad (3.4)$$

Considérons le résultat *(R2)* appliqué à $\langle\langle \cdot, U\rangle\rangle$ et à la constante $c = \max \{ \langle\langle M, U\rangle\rangle \mid M \in \mathfrak{J}F(x_o) \}$.

Selon ce résultat, pour que l'inégalité (3.4) soit vérifiée, il suffit de montrer que

$$\limsup_{n \to \infty} \langle\langle JF(x_n), U\rangle\rangle \leq \max \{ \langle\langle M, U\rangle\rangle \mid M \in \mathfrak{J}F(x_o) \}$$

pour toute suite $(x_n)_{n=1}^{\infty}$ de points de Ω_F^c qui converge vers x_o.

Soit $(x_n)_{n=1}^{\infty}$ une telle suite. Comme l'application JF est localement bornée, la suite $(JF(x_n))_{n=1}^{\infty}$ est bornée. Ainsi, par application du théorème de Bolzano-Weierstrass, il existe une sous-suite $(JF(x_{q_n}))_{n=1}^{\infty}$, extraite de $(JF(x_n))_{n=1}^{\infty}$, qui converge vers une matrice M^* appartenant à $\mathcal{M}(q,p)$.

Cette matrice M^* appartient à $\mathfrak{J}F(x_o)$ car elle vérifie les conditions de définition de $\mathfrak{J}F(x_o)$.

Puisque l'application $\langle\langle \cdot , U \rangle\rangle$ est continue, nous avons que

$$\langle\langle M^*, U \rangle\rangle = \lim_{n \to \infty} \langle\langle JF(x_{q_n}), U \rangle\rangle.$$

En prenant, parmi les matrices M^* qui sont limite d'une sous-suite extraite de $(JF(x_n))_{n=1}^{\infty}$, celle pour laquelle $\langle\langle M^*, U \rangle\rangle$ a la plus grande valeur, nous obtenons que

$$\langle\langle M^*, U \rangle\rangle = \limsup_{n \to \infty} \langle\langle JF(x_n), U \rangle\rangle$$

car la limite supérieure d'une suite est sa plus grande valeur d'adhérence.

Donc, nous obtenons que

$$\limsup_{n \to \infty} \langle\langle JF(x_n), U \rangle\rangle \le \max \{ \langle\langle M, U \rangle\rangle \mid M \in \mathfrak{J}F(x_o) \},$$

ce qui est équivalent, par le résultat *(R2)* à (3. 4).

Les assertions (3. 3) et (3. 4) étant vérifiées, la thèse est démontrée.

□

Notons la matrice U par $[u_1, \ldots, u_q]^T$, où u_i appartient à $\mathbb{R}^p$ pour tout $1 \le i \le q$. Ainsi, la définition du produit scalaire $\langle\langle \cdot , \cdot \rangle\rangle$ permet d'exprimer la proposition 3.2.1 de la façon suivante:

Proposition 3.2.2 : Pour toute matrice U de $\mathcal{M}(q,p)$,

$$\max \{ \langle\langle M, U \rangle\rangle \mid M \in \mathfrak{J}F(x_o) \} = \limsup_{\substack{x \to x_o \\ x \in \Omega_F^c}} \sum_{i=1}^{q} \langle \nabla f_i(x), u_i \rangle.$$

Cette expression n'est pas encore très manipulable. C'est pourquoi nous allons nous intéresser aux matrices U de $\mathcal{M}(q,p)$ pouvant s'écrire comme le produit tensoriel de deux vecteurs u de $\mathbb{R}^p$ et v de $\mathbb{R}^q$, c'est-à-dire telles que $U = u \otimes v$. Ainsi, nous obtenons le résultat suivant.

Proposition 3.2.3 : Pour tout u appartenant à $\mathbb{R}^p$ et v appartenant à $\mathbb{R}^q$,

$$\max \{ \langle Mu, v \rangle \mid M \in \mathfrak{J}F(x_o) \} = \limsup_{\substack{x \to x_o \\ x \in \Omega_F^c}} \langle JF(x)u, v \rangle .$$

Preuve :

Soient u appartenant à $\mathbb{R}^p$, v appartenant à $\mathbb{R}^q$ et A une matrice de $\mathcal{M}(q,p)$.

Remarquons que $\langle\langle A, u\otimes v \rangle\rangle = \langle Au, v \rangle$.

En effet, posons $u = (u_i)_{i=1}^p$, $v = (v_i)_{i=1}^q$ et $A = (a_{ij})$, avec $1 \le i \le q$ et $1 \le j \le p$. Nous calculons que:

$$\langle\langle A, u\otimes v \rangle\rangle = \text{trace } (A \cdot (u\otimes v)^T)$$

$$= \text{trace} \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ a_{q1} & \cdots & a_{qp} \end{pmatrix} \cdot \begin{pmatrix} u_1 v_1 & \cdots & u_p v_1 \\ u_1 v_q & \cdots & u_p v_q \end{pmatrix}^T ,$$

$$= \text{trace} \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ a_{q1} & \cdots & a_{qp} \end{pmatrix} \cdot \begin{pmatrix} u_1 v_1 & \cdots & u_1 v_q \\ u_p v_1 & \cdots & u_p v_q \end{pmatrix} ,$$

$$= \text{trace} \begin{pmatrix} \sum_{k=1}^{q} a_{1k}\, u_k\, v_1 & \cdots \\ \cdots & \sum_{k=1}^{q} a_{qk}\, u_k\, v_q \end{pmatrix} ,$$

$$= \sum_{h=1}^{q} \sum_{k=1}^{p} a_{hk} \, u_k \, v_h \, , \qquad (3.5)$$

et que

$$\langle Au, v\rangle = \left[\begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \\ a_{q1} & \cdots & a_{qp} \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix} \right]^T \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_q \end{pmatrix},$$

$$= \left(\sum_{k=1}^{p} a_{1k} \, u_k \, , \, \ldots \, , \sum_{k=1}^{p} a_{qk} \, u_k \right) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_q \end{pmatrix},$$

$$= \sum_{h=1}^{q} \sum_{k=1}^{p} a_{hk} \, u_k \, v_h \, . \qquad (3.6)$$

Les égalités (3.5) et (3.6) impliquent que $\langle\langle A, u\otimes v\rangle\rangle = \langle Au, v\rangle$.

De cette remarque il suit que

$$\max \{ \langle\langle M, u\otimes v\rangle\rangle \mid M \in \mathfrak{J}F(x_o) \} = \max \{ \langle Mu, v\rangle \mid M \in \mathfrak{J}F(x_o) \},$$

et que

$$\langle\langle JF(x), U\rangle\rangle = \langle JF(x)u, v\rangle,$$

pour tout x appartenant à Ω_F^c.

Ainsi, la proposition 3.2.1 implique que:

$$\max \{ \langle Mu, v\rangle \mid M \in \mathfrak{J}F(x_o) \} = \limsup_{\substack{x \to x_o \\ x \in \Omega_F^c}} \langle JF(x)u, v\rangle.$$

□

Dans cette proposition 3.2.3, en exprimant le membre de droite en termes de différences de quotients des fonctions f_i, nous obtenons le théorème suivant.

Théorème 3.2.4 : Pour tout u appartenant à $\mathbb{R}^p$ et $v = (v_i)_{i=1}^{q}$ appartenant à $\mathbb{R}^q$,

$$\max \{ \langle Mu, v\rangle \mid M \in \Im F(x_o) \} = \limsup_{\substack{x \to x_o \\ \lambda \to 0^+}} \sum_{i=1}^{q} \frac{v_i [f_i(x+\lambda u) - f_i(x)]}{\lambda}.$$

Preuve :

Soit v un point de $\mathbb{R}^q$. Considérons la fonction F_v définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles telle que $F_v(x) = \langle F(x), v\rangle$ pour tout x appartenant à $\mathcal{O}$.

La propriété annexe A.3 de cette fonction F_v implique que F_v est localement Lipschitzienne en tout point de $\mathcal{O}$ car F l'est. Donc, le gradient généralisé de F_v peut être envisagé.

De plus, $\partial F_v(x_o) = [\Im F(x_o)]^T v$. (3.7)

Remarque: La démonstration de cette égalité étant assez longue, au lieu de la développer dans ce chapitre, celle-ci sera expliquée dans le cas particulier A.4.2 de la propriété annexe A.4.

Soit u un point de $\mathbb{R}^p$. La propriété 2.2.4 caractérisant le gradient généralisé par sa fonction de support nous indique que:

$$\max \{ \langle \xi, u\rangle \mid \xi \in \partial F_v(x_o) \} = F_v^{\circ}(x_o;u) , \tag{3.8}$$

où $F_v^{\circ}(x_o;u)$ désigne la dérivée directionnelle généralisée de F_v au point x_o dans la direction u.

En tenant compte de l'égalité (3.7), (3.8) devient

$$\max \{ \langle u, M^T v\rangle \mid M \in \Im F(x_o) \} = F_v^{\circ}(x_o;u),$$

c'est-à-dire que

$$\max \{ \langle Mu, v\rangle \mid M \in \Im F(x_o) \} = F_v^{\circ}(x_o;u). \tag{3.9}$$

Or, par définition de la dérivée directionnelle généralisée, nous avons que $F_v^{\circ}(x_o;u) = \limsup_{\substack{x \to x_o \\ \lambda \to 0^+}} \dfrac{F_v(x+\lambda u) - F_v(x)}{\lambda}$.

L'expression de F_v transforme alors l'égalité (3.9) comme suit

$$\max \{ \langle Mu,v \rangle \mid M \in \mathfrak{J}F(x_o) \} = \limsup_{\substack{x \to x_o \\ \lambda \to 0^+}} \frac{\langle F(x+\lambda u) - F(x),v \rangle}{\lambda},$$

c'est-à-dire que

$$\max \{ \langle Mu,v \rangle \mid M \in \mathfrak{J}F(x_o) \} = \limsup_{\substack{x \to x_o \\ \lambda \to 0^+}} \sum_{i=1}^{q} \frac{v_i [f_i(x+\lambda u) - f_i(x)]}{\lambda}.$$

□

Sachant que n'importe quelle matrice de $\mathcal{M}(q,p)$ peut s'écrire comme $\sum_{k=1}^{r} a_k \otimes b_k$, où $r = \min (q,p)$, et a_k appartient à $\mathbb{R}^p$ et b_k appartient à $\mathbb{R}^q$ pour tout k tel que $1 \le k \le r$ (la preuve de cette propriété se trouve en annexe C.1), il faut pouvoir calculer

$$\max \{ \sum_{k=1}^{r} \langle Ma_k,b_k \rangle \mid M \in \mathfrak{J}F(x_o) \}$$

afin d'exprimer la fonction de support de $\mathfrak{J}F(x_o)$.

Cependant, nous ne connaissons aucune propriété pouvant évaluer cette expression. En effet, le théorème 3.2.4 permet uniquement de caractériser $\max \{ \langle Ma_k,b_k \rangle \mid M \in \mathfrak{J}F(x_o) \}$ pour tout k tel que $1 \le k \le r$.

Cette difficulté à trouver la fonction de support de $\mathfrak{J}F(x_o)$ nous conduit à définir l'objet suivant.

Définition 3.2.5 : La <u>bifonction de support</u> d'un ensemble convexe fermé $\mathcal{K}$ est la fonction $\mathcal{S}$ définie sur $\mathbb{R}^p \times \mathbb{R}^q$ et à valeurs réelles par $\mathcal{S}(u,v) = \max \{ \langle Mu,v\rangle \mid M \in \mathcal{K} \}$ pour tout u appartenant à $\mathbb{R}^p$ et v appartenant à $\mathbb{R}^q$.

La propriété 3.1.2 indique que $\mathfrak{J}F(x_o)$ est un ensemble convexe et fermé. Ainsi, nous pouvons envisager la bifonction de support de $\mathfrak{J}F(x_o)$. Nous la noterons par $F^{\circ}(x_o;\cdot,\cdot)$, et nous obtenons donc que, pour tout u appartenant à $\mathbb{R}^p$ et v appartenant à $\mathbb{R}^q$,

$$F^{\circ}(x_o;u,v) = \max \{ \langle Mu,v\rangle \mid M \in \mathfrak{J}F(x_o) \}. \tag{3.10}$$

De plus, d'après le théorème 3.2.4, nous avons que

$$F^{\circ}(x_o;u,v) = \limsup_{\substack{x \to x_o \\ \lambda \to 0^+}} \sum_{i=1}^{q} \frac{v_i\,[f_i(x+\lambda u) - f_i(x)]}{\lambda}, \tag{3.11}$$

quel que soient u dans $\mathbb{R}^p$ et v dans $\mathbb{R}^q$.

Afin de mieux voir ce que peut représenter cette bifonction de support de $\mathfrak{J}F(x_o)$, nous allons développer deux cas particuliers.

Cas particuliers 3.2.6 :

(C1) Supposons que q = 1, c'est-à-dire que F est une fonction à valeurs réelles.

Soient u un point de $\mathbb{R}^p$ et v un nombre réel.

En tenant compte de (3.11), la bifonction de support de $\mathfrak{J}F(x_o)$ est

$$F^\circ(x_o;u,v) = \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{v\cdot[F(x+\lambda u) - F(x)]}{\lambda}.$$

Si $v \geq 0$, il suit que

$$F^\circ(x_o;u,v) = v \cdot \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{[F(x+\lambda u) - F(x)]}{\lambda}$$

et, par définition de la dérivée directionnelle généralisée de F au point x_o dans la direction u,

$$F^\circ(x_o;u,v) = v \cdot F^\circ(x_o;u).$$

Si $v \leq 0$, comme $(-v)$ est un nombre positif, il suit que

$$F^\circ(x_o;u,v) = (-v) \cdot \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{[F(x) - F(x+\lambda u)]}{\lambda}.$$

Dans l'expression de la limite supérieure, changeons la notation de x par $y = (x+\lambda u)$. Nous pouvons le réaliser car, comme x tend vers x_o et λ tend vers 0 ,le terme $(x+\lambda u)$ converge vers x_o. Ainsi,

$$F^\circ(x_o;u,v) = (-v) \cdot \limsup_{\substack{y \to x_o \\ \lambda \to o^+}} \frac{[F(y-\lambda u) - F(y)]}{\lambda},$$

et la définition de dérivée directionnelle généralisée de F au point x_o dans la direction $(-u)$ implique que

$$F^\circ(x_o;u,v) = (-v) \cdot F^\circ(x_o;-u).$$

En résumé, lorsque F est une fonction à valeurs réelles, la relation entre la bifonction de support de $\mathfrak{J}F(x_o)$ et la fonction de support de $\mathfrak{J}F(x_o)$ est

$$F^{\circ}(x_o;u,v) = \begin{cases} v \cdot F^{\circ}(x_o;u) & \text{si } v \geq 0 \\ (-v) \cdot F^{\circ}(x_o;-u) & \text{si } v \leq 0 \end{cases},$$

quels que soient u dans $\mathbb{R}^p$ et v un nombre réel.

■

(C2) Supposons que p = 1. La fonction F est alors définie sur un intervalle réel ouvert $]\alpha,\beta[$ et à valeurs dans $\mathbb{R}^q$.

Soient u un nombre réel et v un point de $\mathbb{R}^q$.

Remarque : Pour désigner un élément de $\mathfrak{J}F(x_o)$, nous utiliserons la notation "x" car, comme p = 1, un élément de $\mathfrak{J}F(x_o)$ est un vecteur de $\mathbb{R}^q$, c'est-à-dire que les matrices de la définition de $\mathfrak{J}F(x_o)$ ont une colonne.

Par définition, la bifonction de support de $\mathfrak{J}F(x_o)$ est

$$F^{\circ}(x_o;u,v) = \max \{ \langle xu,v \rangle \mid x \in \mathfrak{J}F(x_o) \}.$$

Le nombre u étant réel, il suit que

$$F^{\circ}(x_o;u,v) = \max \{ \langle x,uv \rangle \mid x \in \mathfrak{J}F(x_o) \}. \qquad (3.12)$$

En utilisant cette forme de la bifonction de support de $\mathfrak{J}F(x_o)$, nous allons pouvoir caractériser $\mathfrak{J}F(x_o)$ grâce à la propriété ci-après.

Propriété 3.2.7 : Les deux assertions suivantes sont équivalentes:

i) x^* appartient à $\mathfrak{J}F(x_o)$,

ii) $\langle x^*,v\rangle \le \limsup\limits_{\substack{x \to x_o \\ \lambda \to o^+}} \dfrac{\langle v,F(x+\lambda) - F(x)\rangle}{\lambda}$,

quel que soit v dans $\mathbb{R}^q$.

Preuve :

1°) Montrons que si x^* appartient à $\mathfrak{J}F(x_o)$, (3.13)

$$\text{alors } \langle x^*,v\rangle \le \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle v,F(x+\lambda) - F(x)\rangle}{\lambda},$$

quel que soit v dans $\mathbb{R}^q$.

Soient u un nombre réel et v un point de $\mathbb{R}^q$.

Le théorème 3.2.4 et (3.12) entraînent que

$$\max\ \{\langle x,uv\rangle \mid x \in \mathfrak{J}F(x_o)\} = \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle uv,F(x+\lambda) - F(x)\rangle}{\lambda}. \qquad (3.14)$$

Si $u \ge 0$, (3.14) devient

$$u\cdot\max\ \{\ \langle x,v\rangle \mid x \in \mathfrak{J}F(x_o)\ \} = u\cdot\limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle v,F(x+\lambda) - F(x)\rangle}{\lambda},$$

c'est-à-dire que

$$\max\ \{\ \langle x,v\rangle \mid x \in \mathfrak{J}F(x_o)\ \} = \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle v,F(x+\lambda) - F(x)\rangle}{\lambda}.$$

Si $u \le 0$, (3.14) peut être réécrite comme suit

$$-u\cdot\max\ \{\ \langle x,-v\rangle \mid x \in \mathfrak{J}F(x_o)\ \} = -u\cdot\limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle -v,F(x+\lambda) - F(x)\rangle}{\lambda},$$

c'est-à-dire que

$$\max \{ \langle x,-v\rangle \mid x \in \Im F(x_o) \} = \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle -v, F(x+\lambda) - F(x)\rangle}{\lambda}.$$

Cette dernière égalité étant vérifiée pour tout point v de $\mathbb{R}^q$, elle l'est pour (−v). Ainsi,

$$\max \{ \langle x,v\rangle \mid x \in \Im F(x_o) \} = \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle v, F(x+\lambda) - F(x)\rangle}{\lambda}. \qquad (3.15)$$

Donc, cette égalité (3.15) est vraie quel que soit le nombre réel u.

De plus, par hypothèse, x^* appartient à $\Im F(x_o)$. D'où,

$$\langle x^*,v\rangle \le \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle v, F(x+\lambda) - F(x)\rangle}{\lambda}.$$

2°) Montrons que si $\langle x^*,v\rangle \le \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle v, F(x+\lambda) - F(x)\rangle}{\lambda}$, (3.16)

quel que soit v dans $\mathbb{R}^q$,

alors x^* appartient à $\Im F(x_o)$.

Soit x^* un point de $\mathbb{R}^q$ tel que

$$\langle x^*,v\rangle \le \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle v, F(x+\lambda) - F(x)\rangle}{\lambda},$$

quel que soit v dans $\mathbb{R}^q$.

Supposons que x^* n'appartiennne pas à $\Im F(x_o)$. Ainsi, $\Im F(x_o)$ étant un ensemble convexe compact et $\{x^*\}$ étant un ensemble convexe fermé, le théorème 1.7.2 de séparation stricte des convexes peut être appliqué à ces ensembles.

Donc, il existe un point v^* de $\mathbb{R}^q$ et un nombre réel γ strictement positif tels que

$\langle x,v^* \rangle < \gamma < \langle x^*,v^* \rangle$ pour tout x dans $\mathfrak{J}F(x_o)$.

D'où,

$$\max \{ \langle x,v^* \rangle \mid x \in \mathfrak{J}F(x_o) \} \leq \gamma < \langle x^*,v^* \rangle.$$

Cette inégalité est contraire à l'hypothèse de (3.16). La supposition que x^* n'appartient pas $\mathfrak{J}F(x_o)$ est donc absurde et nous obtenons que x^* appartient à $\mathfrak{J}F(x_o)$.

Les propositions (3.13) et (3.16) étant simultanément vraies, la thèse est démontrée.

□

Ces cas particuliers illustrent la caractérisation de $\mathfrak{J}F(x_o)$ par sa bifonction de support. En effet, pour un élément u de $\mathbb{R}^p$ donné, la fonction $F^\circ(x_o;u,\cdot)$ définie sur $\mathbb{R}^q$ et à valeurs réelles est la fonction de support de $\left[\mathfrak{J}F(x_o)\cdot u\right]$ car,

$$F^\circ(x_o;u,v) = \max \{ \langle Mu,v \rangle \mid M \in \mathfrak{J}F(x_o) \},$$

d'où

$$F^\circ(x_o;u,v) = \max \{ \langle Mu,v \rangle \mid Mu \in \mathfrak{J}F(x_o)\cdot u \},$$

c'est-à-dire

$$F^\circ(x_o;u,v) = \max \{ \langle x,v \rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \}.$$

3.3 INCONVÉNIENTS DE LA DÉFINITION.

Soit x_o un point de $\mathcal{O}$. Pour décrire le premier inconvénient de la définition de $\mathfrak{J}F(x_o)$, nous allons commencer par prouver la propriété qui suit. Dans ce but, nous avons besoin des définitions suivantes.

Soient Λ inclus dans Ω_F^c tel que le complémentaire de Λ dans $\mathcal{O}$ soit de mesure de Lebesgue nulle, et l'ensemble $\mathfrak{J}_\Lambda F(x_o)$ défini de la même manière que $\mathfrak{J}F(x_o)$, mais où les suites de points qui convergent vers x_o sont incluses dans Λ. Autrement dit,

$$\mathfrak{J}_\Lambda F(x_o) = \text{co } [\mathfrak{J}_\Lambda(F, x_o)]$$

$$\text{où } \mathfrak{J}_\Lambda(F, x_o) = \Big\{ M \in \mathcal{M}(q,p) \mid \exists (x_n)_{n=1}^{\infty} \subseteq \Lambda \text{ convergente vers } x_o \text{ telle que } M = \lim_{n\to\infty} JF(x_n) \Big\}. \qquad (3.17)$$

Commençons par montrer la propriété suivante.

Propriété 3.3.1 : Pour tout u appartenant à $\mathbb{R}^p$ et v appartenant à $\mathbb{R}^q$,

$$F_o^\circ(x\ ; u, v) = \max\ \{ \langle Mu, v\rangle \mid M \in \mathfrak{J}_\Lambda F(x_o) \}, \qquad (3.18)$$

$$\text{et, } \mathfrak{J}F(x_o)\cdot u = \mathfrak{J}_\Lambda F(x_o)\cdot u. \qquad (3.19)$$

Preuve :

1°) Remarquons que si (3.19) est vérifié, alors, par définition de $F^\circ(x_o ; u, v)$, (3.18) l'est également.

2°) Soit u un point de $\mathbb{R}^q$. Les ensembles $[\mathfrak{J}F(x_o)\cdot u]$ et $[\mathfrak{J}_\Lambda F(x_o)\cdot u]$ sont des ensembles convexes fermés et compacts.

En effet, la propriété 3.1.2 nous indique que $\mathfrak{J}F(x_o)$ est un ensemble convexe fermé et compact. La même démonstration de cette propriété 3.1.2 appliquée à $\mathfrak{J}_\Lambda F(x_o)$ implique que $\mathfrak{J}_\Lambda F(x_o)$ est également fermé et compact.

Ainsi, pour démontrer que $\mathfrak{J}F(x_o)\cdot u = \mathfrak{J}_\Lambda F(x_o)\cdot u$, il suffit de montrer, par application de la propriété 1.7.4 concernant les ensembles convexes fermés, que la fonction de support de $\left[\mathfrak{J}F(x_o)\cdot u\right]$ est la fonction de support de $\left[\mathfrak{J}_\Lambda F(x_o)\cdot u\right]$.

Il faut donc que,

$$\max \{ \langle x,v\rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \} = \max \{ \langle x,v\rangle \mid x \in \mathfrak{J}_\Lambda F(x_o)\cdot u \},$$

pour tout v dans $\mathbb{R}^q$.

Or, $\max \{ \langle x,v\rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \} = \max \{ \langle Mu,v\rangle \mid M \in \mathfrak{J}F(x_o) \}$.

De plus, la propriété 3.2.3 entraîne que

$$\max \{ \langle Mu,v\rangle \mid M \in \mathfrak{J}F(x_o) \} = \limsup_{\substack{x \to x_o \\ x \in \Omega_F^c}} \langle JF(x)u,v\rangle.$$

D'où, $\max \{ \langle x,v\rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \} = \limsup_{\substack{x \to x_o \\ x \in \Omega_F^c}} \langle u,[JF(x)]^T v\rangle$.

Pour continuer, considérons la fonction F_v définie sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles par $F_v(x) = \langle F(x),v\rangle$. La propriété annexes A.1 de cette fonction F_v implique qu'elle est différentiable en tout point x de $\mathcal{O}$ ou F est différentiable et que

$$\nabla F_v(x) = [JF(x)]^T\cdot v. \tag{3.20}$$

Donc, $\max \{ \langle x,v\rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \} = \limsup_{\substack{x \to x_o \\ x \in \Omega_F^c}} \langle u,\nabla F_v(x)\rangle$.

Par la propriété annexe A.2 de la fonction F_v, nous savons que Ω_{F_v} est inclus dans Ω_F, où Ω_{F_v} est l'ensemble des points x de $\mathcal{O}$ tels que F_v n'est pas différentiable en x.

Ainsi, Ω_F^c est inclus dans $\Omega_{F_v}^c$, d'où $\Omega_F^c = \Omega_{F_v}^c \cap \Omega_F^c$. De cette égalité, il suit que

$$\max \{ \langle x,v \rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \} = \limsup_{\substack{x \to x_o \\ x \in \Omega_{F_v}^c \cap \Omega_F^c}} \langle u,\nabla F_v(x) \rangle.$$

Comme nous travaillons en dimension finie, la caractérisation 2.2.11 appliquée à de la dérivée directionnelle de F_v au point x_o dans la direction u entraîne que

$$\max \{ \langle x,v \rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \} = F_v^{\circ}(x_o;u).$$

Par cette même forme de la dérivée directionnelle de F_v au point x_o dans la direction u, il suit directement que

$$\max \{ \langle x,v \rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \} = \limsup_{\substack{x \to x_o \\ x \in \Lambda \cap \Omega_{F_v}^c}} \langle u,\nabla F_v(x) \rangle,$$

car le complémentaire de Λ dans $\mathcal{O}$ est de mesure de Lebesgue nulle.

En tenant compte de l'égalité (3.20), la dernière égalité devient

$$\max \{ \langle x,v \rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \} = \limsup_{\substack{x \to x_o \\ x \in \Lambda \cap \Omega_{F_v}^c}} \langle u,[JF_v(x)]^T v \rangle.$$

Or, $\Lambda = \Lambda \cap \Omega_{F_v}^c$, donc

$$\max \{ \langle x,v \rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \} = \limsup_{\substack{x \to x_o \\ x \in \Lambda}} \langle u,[JF_v(x)]^T v \rangle.$$

Sans difficulté supplémentaire dans la démonstration, nous pouvons généraliser la propriété 3.2.3 à $\mathfrak{J}_\Lambda F(x_o)$.
Ce qui conduit à l'égalité qui suit

$$\max \{ \langle x,v\rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \} = \max \{ \langle Mu,v\rangle \mid M \in \mathfrak{J}_\Lambda F(x_o) \},$$

c'est-à-dire

$$\max \{ \langle x,v\rangle \mid x \in \mathfrak{J}F(x_o)\cdot u \} = \max \{ \langle x,v\rangle \mid x \in \mathfrak{J}_\Lambda F(x_o)\cdot u \}.$$

Finalement, grâce à la propriété 1.7.4, nous obtenons que

$$\mathfrak{J}F(x_o)\cdot u = \mathfrak{J}_\Lambda F(x_o)\cdot u.$$

□

Selon cette propriété, les ensembles $\left[\mathfrak{J}F(x_o)\cdot u\right]$ et $\left[\mathfrak{J}_\Lambda F(x_o)\cdot u\right]$ sont égaux quels que soient Λ inclus dans Ω_F^c dont le complémentaire dans $\mathcal{O}$ est de mesure de Lebesgue nulle et u dans $\mathbb{R}^p$. Cependant, rien ne permet d'affirmer que les ensembles $\mathfrak{J}F(x_o)$ et $\mathfrak{J}_\Lambda F(x_o)$ sont identiques, à l'exception de deux cas particuliers (Nous développerons ces cas particuliers un peu plus loin). Donc, la définition de $\mathfrak{J}F(x_o)$ ne tient pas compte de tous les ensembles dont le complémentaire dans $\mathcal{O}$ est de mesure de Lebesgue nulle, contrairement au gradient généralisé comme l'indique la propriété 2.2.10.

Envisageons les cas particuliers.

Cas particuliers 3.3.2 :

(C1) Supposons que q = 1, c'est-à-dire que F est une fonction à valeurs réelles.

La caractérisation 2.2.10 du gradient généralisé indique que, pour tout ensemble S de mesure de Lebesgue nulle,

$$\partial F(x_o) = \{ x \in \mathbb{R}^p \mid \exists (x_n)_{n=1}^{\infty} \subseteq S^c \cap \Omega_F^c \text{ convergente vers } x_o \text{ telle que } x = \lim_{n\to\infty} \nabla F(x_n) \}.$$

En prenant $S = \Omega_F$, respectivement $S = \Lambda^c$, nous obtenons que $\mathfrak{J}F(x_o) = [\partial F(x_o)]^T$, respectivement $\mathfrak{J}_\Lambda F(x_o) = [\partial F(x_o)]^T$.
Donc $\mathfrak{J}F(x_o) = \mathfrak{J}_\Lambda F(x_o)$.

■

(C2) Supposons que p = 1.
Par la propriété 3.3.1, nous savons que $\mathfrak{J}F(x_o)\cdot u = \mathfrak{J}_\Lambda F(x_o)\cdot u$ pour tout nombre réel u. Donc $\mathfrak{J}F(x_o) = \mathfrak{J}_\Lambda F(x_o)$.

■

La définition de $\mathfrak{J}F(x_o)$ a un autre désavantage. Bien que grâce à la définition de la bifonction de support de $\mathfrak{J}F(x_o)$, toutes les matrices M de $\mathfrak{J}F(x_o)$ satisfont l'inégalité

$$\langle Mu, v\rangle \leq F^{\circ}(x_o; u, v)$$

pour tout u appartenant à $\mathbb{R}^p$ et v appartenant à $\mathbb{R}^q$, nous ne pouvons répondre à la question suivante.
Quelles sont les matrices M de $\mathcal{M}(q,p)$ telles que $\langle Mu, v\rangle \leq F^{\circ}(x_o; u, v)$ pour tout u dans $\mathbb{R}^p$ et v dans $\mathbb{R}^q$?
Autrement dit, quelles sont les matrices M de $\mathcal{M}(q,p)$ telles que $M\cdot u$ appartient à $\left[\mathfrak{J}F(x_o)\cdot u\right]$ pour tout point u de $\mathbb{R}^p$?

La prochaine propriété que nous allons développer montre l'équivalence entre ces deux questions.

Propriété 3.3.3 : Pour toute matrice M de $\mathcal{M}(q,p)$, les assertions suivantes sont équivalentes.

i) Pour tout u appartenant à $\mathbb{R}^p$ et v appartenant à $\mathbb{R}^q$,

$$\langle Mu,v\rangle \le F^{\circ}(x_o;u,v). \tag{3.21}$$

ii) Pour tout u appartenant à $\mathbb{R}^p$,

$$M\cdot u \text{ appartient à } \left[\mathfrak{J}F(x_o)\cdot u\right]. \tag{3.22}$$

Preuve :

Soit M une matrice de $\mathcal{M}(q,p)$.

1°) Montrons que si $\langle Mu,v\rangle \le F^{\circ}(x_o;u,v)$ quels que soient les points u de $\mathbb{R}^p$ et v de $\mathbb{R}^q$,

alors $M\cdot u$ appartient à $\left[\mathfrak{J}F(x_o)\cdot u\right]$ pour tout point u de $\mathbb{R}^p$.

Soient les points u de $\mathbb{R}^p$ et v de $\mathbb{R}^q$. Considérons l'ensemble auxiliaire $\mathcal{Z}(x_o;u,v)$ des matrices M de $\mathcal{M}(q,p)$ telles que

$\langle Mu,v\rangle \le F^{\circ}(x_o;u,v)$.

Autrement dit,

$$\mathcal{Z}(x_o;u,v) = \left\{ M \in \mathcal{M}(q,p) \mid \langle Mu,v\rangle \le F^{\circ}(x_o;u,v) \right\}.$$

1.1°) $\mathcal{Z}(x_o;u,v)$ est un ensemble convexe.

En effet, soient M et N dans $\mathcal{Z}(x_o;u,v)$ et soit λ un nombre réel compris entre 0 et 1, nous savons que

$\lambda\cdot\langle Mu,v\rangle + (1-\lambda)\cdot\langle Nu,v\rangle \le F^{\circ}(x_o;u,v)$.

Ainsi, la linéarité du produit scalaire par rapport à la première variable implique que

$\langle(\lambda\cdot M + (1-\lambda)\cdot N)\cdot u,v\rangle \le F^{\circ}(x_o;u,v)$.

Et donc, $(\lambda \cdot M + (1-\lambda) \cdot N)$ appartient à $\mathcal{Z}(x_o;u,v)$.

1.2°) $\mathcal{Z}(x_o;u,v)$ est un ensemble fermé.

Pour montrer cette caractéristique de $\mathcal{Z}(x_o;u,v)$, il faut que toute suite $(M_n)_{n=1}^{\infty}$ de matrices de $\mathcal{Z}(x_o;u,v)$ qui converge vers une matrice M de $\mathcal{M}(q,p)$ soit telle que cette matrice M appartient à $\mathcal{Z}(x_o;u,v)$.

Pour commencer, remarquons que, par continuité du produit scalaire, la suite $(\langle M_n u,v\rangle)_{n=1}^{\infty}$ converge vers $\langle Mu,v\rangle$. D'où, comme $\langle M_n u,v\rangle \leq F^{\circ}(x_o;u,v)$ pour tout $n \geq 1$, il suit par passage à la limite que $\langle Mu,v\rangle \leq F^{\circ}(x_o;u,v)$. Donc M appartient à $\mathcal{Z}(x_o;u,v)$.

Finalement, $\mathcal{Z}(x_o;u,v)$ est un ensemble convexe fermé.

1.3°) Montrons la thèse.

Par construction de $\mathcal{Z}(x_o;u,v)$, $\langle Mu,v\rangle \leq F^{\circ}(x_o;u,v)$ quel que soit M dans $\mathcal{Z}(x_o;u,v)$. D'où,

$$\sup \{ \langle Mu,v\rangle \mid M \in \mathcal{Z}(x_o;u,v) \} \leq F^{\circ}(x_o;u,v).$$

C'est-à-dire, par définition de la bifonction de support de $\mathfrak{J}F(x_o)$, que

$$\sup \{ \langle Mu,v\rangle \mid M \in \mathcal{Z}(x_o;u,v) \} \leq \max \{ \langle Mu,v\rangle \mid M \in \mathfrak{J}F(x_o) \}.$$

Cette inégalité étant vérifiée quel que soit v dans $\mathbb{R}^q$, la propriété 1.7.4 entraîne que, pour tout u appartenant à $\mathbb{R}^p$,

$$\left[\mathcal{Z}(x_o;u,v)\cdot u\right] \text{ est inclus dans } \left[\mathfrak{J}F(x_o)\cdot u\right],$$

car $\mathcal{Z}(x_o;u,v)$ et $\mathfrak{J}F(x_o)$ sont des ensembles convexes fermés.

Donc, pour toute matrice M de $\mathcal{M}(q,p)$ telles que

$$\langle Mu,v\rangle \leq F^{\circ}(x_o;u,v),$$

nous obtenons que Mu appartient à $\left[\mathfrak{J}F(x_o)\cdot u\right]$.

2°) Montrons que si Mu appartient à $[\mathfrak{J}F(x_o)\cdot u]$ pour tout point u de $\mathbb{R}^p$,

alors $\langle Mu,v\rangle \leq F^{\circ}(x_o;u,v)$ quels que soient les points u de $\mathbb{R}^p$ et v de $\mathbb{R}^q$.

Soit u dans $\mathbb{R}^p$ et soit M une matrice de $\mathcal{M}(q,p)$ tels que Mu appartient à $[\mathfrak{J}F(x_o)\cdot u]$.

Ainsi, il existe une matrice M^* de $\mathfrak{J}F(x_o)$ tels que $Mu = M^*u$. D'où, $\langle Mu,v\rangle = \langle M^*u,v\rangle$ pour tout v appartenant à $\mathbb{R}^q$. Donc,

$$\langle Mu,v\rangle \leq \max \{ \langle Mu,v\rangle \mid M \in \mathfrak{J}F(x_o) \},$$

quels que soient les points u de $\mathbb{R}^p$ et v de $\mathbb{R}^q$. C'est-à-dire, par définition de la bifonction de support de $\mathfrak{J}F(x_o)$, que

$$\langle Mu,v\rangle \leq F^{\circ}(x_o;u,v),$$

quels que soient les points u de $\mathbb{R}^p$ et v de $\mathbb{R}^q$.

□

Comme les propriétés vues jusqu'içi ne nous permettent pas de répondre à la question posée ci-dessus, nous ne pouvons pas déterminer entièrement $\mathfrak{J}F(x_o)$ grâce à sa bifonction de support.

Ces deux inconvénients de la définition de la matrice Jacobienne généralisée en un point x_o n'impliquent pas que cette définition est à rejeter. D'ailleurs, cette définition sera encore utilisée par la suite. Mais, il est important de savoir que cette définition de $\mathfrak{J}F(x_o)$ n'est pas nécessairement ultime.

3.4 ENVELOPPE PLEINE.

Nous allons chercher à développer un objet conservant les propriétés de la matrice Jacobienne généralisée, mais qui n'en a pas les désavantages.

Dans ce but, considérons la définition suivante.

Définition 3.4.1 : Tout <u>sous-ensemble</u> $\mathcal{A}$ de $\mathcal{M}(q,p)$ est dit <u>plein</u> si $\mathcal{A}$ contient toutes les matrices M de $\mathcal{M}(q,p)$ telles que $(M\cdot u)$ appartient à $(\mathcal{A}\cdot u)$ quel que soit u dans $\mathbb{R}^p$.

Pour de tels sous-ensemble de $\mathcal{M}(q,p)$, nous avons la propriété qui suit.

Propriété 3.4.2 : L'intersection d'ensembles pleins est pleine.

Preuve :

Soit $\mathfrak{P}$, l'ensemble des parties pleines de $\mathcal{M}(q,p)$.

Montrons que $\mathcal{K} = \{ \cap \mathcal{A} \mid \mathcal{A} \in \mathfrak{P} \}$ est plein, c'est-à-dire que $\mathcal{K}$ contient toutes les matrices M de $\mathcal{M}(q,p)$ telles que $M\cdot u$ appartient à $\mathcal{K}\cdot u$ pour tout point u de $\mathbb{R}^p$.

Soit M une matrice de $\mathcal{M}(q,p)$ et u un point de $\mathbb{R}^p$ tels que $M\cdot u$ appartient à $\mathcal{K}\cdot u$. Ainsi,

$M\cdot u$ appartient à $\{ (\cap \mathcal{A})\cdot u \mid \mathcal{A} \in \mathfrak{P} \}$.

D'où, $M \cdot u$ appartient à $\mathcal{A} \cdot u$ quel que soit $\mathcal{A}$ dans $\mathfrak{P}$. Comme chaque élément de $\mathfrak{P}$ est plein, il suit que M appartient à $\mathcal{A}$ pour tout $\mathcal{A}$ dans $\mathfrak{P}$, et donc M appartient à $\mathcal{K}$.

□

Cette propriété motive la prochaine définition.

Définition 3.4.3 : L'<u>enveloppe</u> <u>pleine</u> d'un sous-ensemble $\mathcal{X}$ de $\mathcal{M}(q,p)$, notée *plen* $\mathcal{X}$, est le plus petit sous-ensemble plein de $\mathcal{M}(q,p)$ contenant $\mathcal{X}$.

Autrement dit, en notant,

$$\mathfrak{P}(\mathcal{X}) = \{ \mathcal{A} \in \mathcal{M}(q,p) \mid \mathcal{A} \text{ est plein et } \mathcal{X} \subseteq \mathcal{A} \},$$

nous avons que *plen* $\mathcal{X} = \{ \cap \mathcal{A} \mid \mathcal{A} \in \mathfrak{P}(\mathcal{X}) \}$.

Compte tenu de la définition de l'enveloppe pleine d'un sous-ensemble $\mathcal{X}$ de $\mathcal{M}(q,p)$, *plen* $\mathcal{X}$ peut être caractérisé par le théorème suivant.

Théorème 3.4.4 : Soient $\mathcal{X}$ inclus dans $\mathcal{M}(q,p)$ et $\mathcal{W}$ l'ensemble des matrices M de $\mathcal{M}(q,p)$ telles que $M \cdot u$ appartient à $\mathcal{X} \cdot u$ quel que soit u dans $\mathbb{R}^p$, alors *plen* $\mathcal{X} = \mathcal{W}$.

Preuve :

Il faut démontrer que *plen* $\mathcal{X} = \mathcal{W}$, où

$$\mathcal{W} = \{ M \in \mathcal{M}(q,p) \mid M \cdot u \in \mathcal{X} \cdot u \ \forall u \in \mathbb{R}^p \}.$$

Pour commencer, nous allons montrer que $\mathcal{W}$ est un sous-ensemble plein de $\mathcal{M}(q,p)$. Ensuite, nous verrons que $\mathcal{W}$ contient $\mathcal{X}$. Et enfin, nous montrerons que $\mathcal{W}$ est le plus petit sous-ensemble plein de $\mathcal{M}(q,p)$ contenant $\mathcal{X}$.

1°) Montrons que $\mathcal{W}$ est un sous-ensemble plein de $\mathcal{M}(q,p)$.

Soit M une matrice de $\mathcal{M}(q,p)$ tel que $M \cdot u$ appartient à $\mathcal{W} \cdot u$ pour tout point u de $\mathbb{R}^p$. Il faut obtenir que M appartient à $\mathcal{W}$.

Soit u un point de $\mathbb{R}^p$. Nous savons que $M \cdot u$ appartient à $\mathcal{W} \cdot u$. D'où, il existe une matrice N de $\mathcal{W}$ tel que $M \cdot u = N \cdot u$. Or, $N \cdot u$ appartient à $\mathcal{X} \cdot u$. Ainsi, $M \cdot u$ appartient à $\mathcal{X} \cdot u$, c'est-à-dire, par définition de l'ensemble $\mathcal{W}$, que M appartient à $\mathcal{W}$. Donc, $\mathcal{W}$ est un sous-ensemble plein de $\mathcal{M}(q,p)$.

2°) Montrons que $\mathcal{X}$ est inclus dans $\mathcal{W}$.

Cette inclusion est évidente. En effet, soit M une matrice de $\mathcal{X}$, nous avons que $M \cdot u$ appartient à $\mathcal{X} \cdot u$ quel que soit u dans $\mathbb{R}^p$. Donc, M appartient à $\mathcal{W}$ par construction de l'ensemble $\mathcal{W}$.

3°) Montrons que tout sous-ensemble plein de $\mathcal{M}(q,p)$ contenant $\mathcal{X}$ contient également $\mathcal{W}$.

Soit $\mathcal{V}$ une partie pleine de $\mathcal{M}(q,p)$ tel que $\mathcal{X}$ soit inclus dans $\mathcal{V}$, nous devons obtenir que $\mathcal{W}$ est inclus dans $\mathcal{V}$.

Soit M une matrice de $\mathcal{W}$. D'où, $M \cdot u$ appartient à $\mathcal{X} \cdot u$ quel que soit u dans $\mathbb{R}^p$.

Supposons que M n'appartienne pas à $\mathcal{V}$, alors il existe un point u^* de $\mathbb{R}^p$ tel $M \cdot u^*$ n'appartient pas à $\mathcal{V} \cdot u^*$ car $\mathcal{V}$ est un sous-ensemble plein de $\mathcal{M}(q,p)$.

Comme $\mathcal{X}$ est inclus dans $\mathcal{V}$, il suit que $M \cdot u^*$ n'appartient pas à $\mathcal{X} \cdot u^*$, ce qui est contredit l'hypothèse que M appartient à $\mathcal{W}$. Donc, M appartient à $\mathcal{V}$. D'où $\mathcal{W}$ est inclus dans $\mathcal{V}$.

En conclusion $\mathcal{W}$ est le plus petit sous-ensemble plein de $\mathcal{M}(q,p)$ contenant $\mathcal{X}$, c'est-à-dire que $\mathit{plen}\ \mathcal{X} = \mathcal{W}$.

□

La propriété qui suit sera utilisée par la suite.

Propriété 3.4.5 : Soit $\mathcal{X}$ un ensemble convexe compact de $\mathcal{M}(q,p)$, alors $\mathit{plen}\ \mathcal{X}$ est un ensemble convexe compact de $\mathcal{M}(q,p)$.

Preuve :

1°) Montrons que $\mathit{plen}\ \mathcal{X}$ est un ensemble convexe de $\mathcal{M}(q,p)$.

Soient M et N dans $\mathit{plen}\ \mathcal{X}$, il faut montrer que $(\lambda \cdot M + (1-\lambda) \cdot N)$ appartient à $\mathit{plen}\ \mathcal{X}$ pour tout nombre λ de l'intervalle $[0,1]$.

La caractérisation 3.4.4 indique que

$$\mathit{plen}\ \mathcal{X} = \{ M \in \mathcal{M}(q,p) \mid M \cdot u \in \mathcal{X} \cdot u,\ \forall u \in \mathbb{R}^p \}.$$

Ainsi, $M \cdot u$ appartient à $\mathcal{X} \cdot u$ et $N \cdot u$ appartient à $\mathcal{X} \cdot u$ quel que soit u de $\mathbb{R}^p$. Soit λ un nombre de l'intervalle $[0,1]$. Comme $\mathcal{X}$ est un ensemble convexe, il suit que $\mathcal{X} \cdot u$ est un ensemble convexe. D'où, $(\lambda \cdot M + (1-\lambda) \cdot N) \cdot u$ appartient à $\mathcal{X} \cdot u$. L'ensemble $\mathit{plen}\ \mathcal{X}$ étant plein par définition, nous obtenons que $(\lambda \cdot M + (1-\lambda) \cdot N)$ appartient à $\mathit{plen}\ \mathcal{X}$. Donc, $\mathit{plen}\ \mathcal{X}$ est convexe.

2°) Montrons que $plen\ \mathcal{X}$ est compact.

Il suffit de prouver que $plen\ \mathcal{X}$ est fermé et borné car $\mathcal{M}(q,p)$ est un espace de dimension finie.

2.1°) Montrons que $plen\ \mathcal{X}$ est fermé.

Considérons la suite $(M_i)_{i=1}^{\infty}$ incluse dans $plen\ \mathcal{X}$ qui converge vers une matrice M de $\mathcal{M}(q,p)$. Il faut voir que M appartient à $plen\ \mathcal{X}$.

Soit i un nombre de $\mathbb{N}_0$. Comme M_i appartient à $plen\ \mathcal{X}$, la caractérisation 3.4.4 entraîne que $M_i \cdot u$ appartient à $\mathcal{X} \cdot u$ pour tout point u de $\mathbb{R}^p$. De plus, nous savons que la suite $(M_i \cdot u)_{i=1}^{\infty}$ converge vers $M \cdot u$. Or, $\mathcal{X} \cdot u$ est un ensemble fermé car $\mathcal{X}$ l'est. D'où, $M \cdot u$ appartient à $\mathcal{X} \cdot u$, c'est-à-dire, par définition de $plen\ \mathcal{X}$, que M appartient à $plen\ \mathcal{X}$.

2.2°) Montrons que $plen\ \mathcal{X}$ est borné, c'est-à-dire qu'il existe $K > 0$ tel que $\|M\| \leq K$ quel que soit M dans $plen\ \mathcal{X}$.

Soit M une matrice de $plen\ \mathcal{X}$. Ainsi, $M \cdot u$ appartient à $\mathcal{X} \cdot u$ pour tout point u de $\mathbb{R}^p$. Comme $\mathcal{X}$ est borné, $\mathcal{X} \cdot u$ l'est également. D'où, pour tout u dans $\mathbb{R}^p$, il existe $K' > 0$ tel que $\|M \cdot u\| \leq K'$.
C'est-à-dire que, pour tout point u de $\mathbb{R}^p$,

$$\sup \{ \|M \cdot u\| \mid M \in plen\ \mathcal{X} \} \leq K'.$$

Le principe 1.9.2 de la borne uniforme, appliqué aux fonctions linéaires définies par les matrices M de $plen\ \mathcal{X}$, implique qu'il existe une constante $K > 0$ telle que

$$\sup \{ \|M\| \mid M \in plen\ \mathcal{X} \} \leq K.$$

Donc $plen\ \mathcal{X}$ est un ensemble borné

□

Soit x_o un point de $\mathcal{O}$.

Nous allons montrer que $\mathit{plen}\ \mathfrak{J}F(x_o)$ a les caractéristiques que nous recherchons. De plus, le théorème que nous venons de démontrer indique que

$$\mathit{plen}\ \mathfrak{J}F(x_o) = \{\ M \in \mathcal{M}(q,p)\ \mid\ M\cdot u \in \mathfrak{J}F(x_o)\cdot u,\ \forall u \in \mathbb{R}^p\ \}.$$

Avant de commencer l'étude de $\mathit{plen}\ \mathfrak{J}F(x_o)$, notons que l'enveloppe pleine ne modifie pas la définition du gradient généralisé en un point x_o d'une fonction à valeurs réelles localement Lipschitzienne. C'est l'objet de la propriété suivante.

Propriété 3.4.6 : Si $q = 1$, alors $\mathit{plen}\ \partial F(x_o) = \partial F(x_o)$.

Preuve :

Pour commencer, remarquons que si $q = 1$, au lieu d'utiliser $\mathfrak{J}F(x_o)$, nous pouvons utiliser $\partial F(x_o)$ car dans ce cas $\mathfrak{J}F(x_o) = [\partial F(x_o))]^T$.

Il suffit de montrer que $\partial F(x_o)$ est plein car, par définition de l'enveloppe pleine d'un sous-ensemble de $\mathbb{R}^p$, $\mathit{plen}\ \partial F(x_o)$ est le plus petit ensemble plein contenant $\partial F(x_o)$.

Donc, nous devons prouver que, si x appartenant à $\mathbb{R}^p$ est tel que $x^T\cdot u = \left([\partial F(x_o)]^T\cdot u\right)$ pour tout u dans $\mathbb{R}^p$, alors x appartient à $\partial F(x_o)$.

Soient x un point de $\mathbb{R}^p$ et u dans $\mathbb{R}^p$ tels que nous ayons

$$x^T\cdot u = \left([\partial F(x_o)]^T\cdot u\right),$$

et soit v dans $\mathbb{R}^p$.

Par définition de la bifonction de support de $\partial F(x_o)$, nous savons que

$$(x^T \cdot u)\cdot v \leq F^{\circ}(x_o;u,v). \tag{3.23}$$

De plus, le premier des cas particuliers 3.2.6 indique que

$$F^{\circ}(x_o;u,v) = \begin{cases} v\cdot F^{\circ}(x_o;u) & \text{si } v \geq 0 \\ (-v)\cdot F^{\circ}(x_o;-u) & \text{si } v \leq 0 \end{cases},$$

D'où, si $v \geq 0$, (3.23) devient

$$(x^T \cdot u)\cdot v \leq v\cdot F^{\circ}(x_o;u).$$

Et comme v est un nombre réel positif, il suit que

$$x^T \cdot u \leq F^{\circ}(x_o;u),$$

c'est-à-dire, par définition de $\partial F(x_o)$, que x appartient à $\partial F(x_o)$ car u a été pris dans $\mathbb{R}^p$ arbitrairement.

De plus, si $v \leq 0$, (3.23) devient

$$(x^T \cdot u)\cdot v \leq (-v)\cdot F^{\circ}(x_o;-u).$$

Et comme $(-v)$ est un nombre réel positif, il suit que

$$x^T \cdot (-u) \leq F^{\circ}(x_o;-u),$$

c'est-à-dire, par définition de $\partial F(x_o)$, que x appartient à $\partial F(x_o)$ car u a été pris dans $\mathbb{R}^p$ arbitrairement. Donc, $\partial F(x_o)$ est un ensemble plein, d'où $\mathit{plen}\ \partial F(x_o) = \partial F(x_o)$.

□

A présent, décrivons les propriétés de $\mathit{plen}\ \mathfrak{J}F(x_o)$. Nous allons commencer par une propriété qui permettra de reformuler plusieurs théorèmes déjà vus.

Propriété 3.4.7 : Pour tout u dans $\mathbb{R}^p$,

$$\mathfrak{J}F(x_o)\cdot u = \mathit{plen}\ \mathfrak{J}F(x_o)\cdot u.$$

Preuve :

Cette propriété suit immédiatement grâce à la caractérisation 3.4.4, c'est-à-dire car

$$\mathit{plen}\ \mathfrak{J}F(x_o) = \{\ M \in \mathcal{M}(q,p)\ |\ M\cdot u \in \mathfrak{J}F(x_o)\cdot u,\ \forall u \in \mathbb{R}^p\ \}.$$

□

Ainsi, suite à cette dernière propriété, nous pouvons réécrire le théorème 3.2.4.

Théorème 3.4.8 : Pour tout u appartenant à $\mathbb{R}^p$ et v appartenant à $\mathbb{R}^q$,

$$F^{\circ}(x_o;u,v) = \max\ \{\ \langle Mu,v\rangle\ |\ M \in \mathit{plen}\ \mathfrak{J}F(x_o)\ \}.$$

De cette propriété, nous déduisons le théorème suivant.

Théorème 3.4.9 : Les deux assertions suivantes sont équivalentes.

i) M appartient à $\mathit{plen}\ \mathfrak{J}F(x_o)$, (3.23)

ii) $\langle Mu,v\rangle \le F^{\circ}(x_o;u,v)$ pour tout u de $\mathbb{R}^p$ et v de $\mathbb{R}^q$. (3.24)

Preuve :

1°) Montrons que si M appartient à $\mathit{plen}\ \mathfrak{J}F(x_o)$,

alors $\langle Mu,v\rangle \le F^{\circ}(x_o;u,v)$ pour tout u dans $\mathbb{R}^p$ et v dans $\mathbb{R}^q$.

Compte tenu de la propriété 3.4.8, cette implication est vérifiée.

2°) Montrons que si $\langle Mu,v\rangle \le F^{\circ}(x_o;u,v)$ pour tout u dans $\mathbb{R}^p$ et v dans $\mathbb{R}^q$,

alors M appartient à *plen* $\mathfrak{J}F(x_o)$.

Soit M une matrice de $\mathcal{M}(q,p)$ telle que $\langle Mu,v\rangle \le F^{\circ}(x_o;u,v)$ pour tout u dans $\mathbb{R}^p$ et v dans $\mathbb{R}^q$.

D'où, par l'équivalence de la propriété 3.3.3, nous savons que

$M\cdot u$ appartient à $\left[\mathfrak{J}F(x_o)\cdot u\right]$.

Or, *plen* $\mathfrak{J}F(x_o) = \{ M \in \mathcal{M}(q,p) \mid M\cdot u \in \mathfrak{J}F(x_o)\cdot u,\ \forall u \in \mathbb{R}^p \}$.

Donc, M appartient à *plen* $\mathfrak{J}F(x_o)$.

□

Maintenant, considérons les objets Λ et $\mathfrak{J}_\Lambda F(x_o)$ définis dans le paragraphe 3.3. Ainsi, la propriété 3.3.1 peut également être reformulée.

Propriété 3.4.10 : *plen* $\mathfrak{J}_\Lambda F(x_o)$ = *plen* $\mathfrak{J}F(x_o)$.

Preuve :

La caractérisation 3.4.4 des ensembles pleins indique que

plen $\mathfrak{J}_\Lambda F(x_o) = \{ M \in \mathcal{M}(q,p) \mid M\cdot u \in \mathfrak{J}_\Lambda F(x_o)\cdot u,\ \forall u \in \mathbb{R}^p \}$.

Or, par la propriété 3.3.1, nous savons que $\mathfrak{J}_\Lambda F(x_o)\cdot u = \mathfrak{J}F(x_o)\cdot u$ quel que soit u appartenant à $\mathbb{R}^p$.

Ainsi,

plen $\mathfrak{J}_\Lambda F(x_o) = \{ M \in \mathcal{M}(q,p) \mid M\cdot u \in \mathfrak{J}F(x_o)\cdot u,\ \forall u \in \mathbb{R}^p \}$.

C'est-à-dire, par la caractérisation 3.4.4, que

$$\mathit{plen}\ \mathfrak{J}_\Lambda F(x_o) = \mathit{plen}\ \mathfrak{J}F(x_o).$$

□

Ces deux dernières propriétés montrent que les inconvénients de la définition de $\mathfrak{J}F(x_o)$ que nous avions cités n'apparaissent plus dans $\mathit{plen}\ \mathfrak{J}F(x_o)$.

En effet, le théorème 3.4.9 permet de caractériser entièrement les matrices M de $\mathcal{M}(q,p)$ telles que

$\langle Mu,v\rangle \le F^{\circ}(x_o;u,v)$ quels que soient u dans $\mathbb{R}^p$ et v dans $\mathbb{R}^q$.

Et grâce à la propriété 3.4.10, les ensembles de mesure de Lebesgue nulle sont pris en compte par la définition de $\mathit{plen}\ \mathfrak{J}F(x_o)$.

Soit v un point de $\mathbb{R}^q$.

Dans la nouvelle propriété de $\mathit{plen}\ \mathfrak{J}F(x_o)$ que nous allons développer, nous utiliserons la fonction F_v définie sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles par $F_v(x) = \langle F(x),v\rangle$ pour tout x dans $\mathcal{O}$.

Propriété 3.4.11 : $\left[\mathit{plen}\ \mathfrak{J}F(x_o)\right]^T$ est le plus grand des ensembles convexes compacts $\mathcal{K}(F;x_o)$ de $\mathcal{M}(q,p)$ pour lesquels

$$\partial F_v(x_o) = \mathcal{K}(F;x_o)\cdot v \qquad (3.25)$$

quel que soit v appartenant à $\mathbb{R}^q$.

Preuve :

Notons que grâce, à la propriété 3.4.5, $\mathit{plen}\ \mathfrak{J}F(x_o)$ est un ensemble convexe compact car $\mathfrak{J}F(x_o)$ l'est.

Soit v dans $\mathbb{R}^q$.

Nous allons commencer par montrer que $\left[\mathit{plen}\ \mathfrak{J}F(x_o)\right]^T = \mathit{plen}\ \left[\mathfrak{J}F(x_o)\right]^T$.
Ensuite, nous remarquerons que $\partial F_v(x_o) = \left[\mathit{plen}\ \mathfrak{J}F(x_o)\right]^T \cdot v$. Enfin, nous verrons que $\left[\mathit{plen}\ \mathfrak{J}F(x_o)\right]^T$ est le plus grand ensemble convexe compact $\mathcal{K}(F;x_o)$ de $\mathcal{M}(q,p)$ vérifiant $\partial F_v(x_o) = \mathcal{K}(F;x_o)\cdot v$.

1°) Montrons que $\left[\mathit{plen}\ \mathfrak{J}F(x_o)\right]^T = \mathit{plen}\ \left[\mathfrak{J}F(x_o)\right]^T$.

Il faut que M appartienne à $\mathit{plen}\ \mathfrak{J}F(x_o)$ soit équivalent à M^T appartient à $\mathit{plen}\ \left[\mathfrak{J}F(x_o)\right]^T$.

Soit M une matrice de $\mathit{plen}\ \mathfrak{J}F(x_o)$. Suite à l'équivalence du théorème 3.4.9, nous obtenons que

$\langle Mu,w\rangle \leq F^{\circ}(x_o;u,w)$ pour tout u dans $\mathbb{R}^p$ et w dans $\mathbb{R}^q$.

C'est-à-dire que

$\langle u,M^Tw\rangle \leq F^{\circ}(x_o;u,w)$ pour tout u dans $\mathbb{R}^p$ et w dans $\mathbb{R}^q$.

Considérons l'élément v de $\mathbb{R}^q$ pris plus haut. Nous savons, par définition de la fonction de support d'un ensemble, que $F^{\circ}(x_o;\cdot,v)$ est la fonction de support de $\partial F_v(x_o)$. Ainsi, $M^T\cdot v$ appartient à $\partial F(x_o)$. Or le cas particulier de la propriété A.4.2 de la propriété annexe A.4 indique que $\partial F_v(x_o) = \left[\mathfrak{J}F(x_o)\right]^T\cdot v$. D'où, $M^T\cdot v$ appartient à $\left[\mathfrak{J}F(x_o)\right]^T\cdot v$. Mais nous pouvons montrer que $\left[\mathfrak{J}F(x_o)\right]^T\cdot v = \left[\mathit{plen}\ \left[\mathfrak{J}F(x_o)\right]^T\right]\cdot v$ (la preuve de égalité vient directement de la caractérisation de l'enveloppe pleine d'un ensemble de $\mathcal{M}(p,q)$). Donc $M^T\cdot v$ appartient à $\left[\mathit{plen}\ \left[\mathfrak{J}F(x_o)\right]^T\right]\cdot v$.

Comme $\mathit{plen}\left[\mathfrak{J}F(x_o)\right]^T$ est un ensemble plein, nous obtenons que M^T appartient à $\mathit{plen}\left[\mathfrak{J}F(x_o)\right]^T$. Donc, $\left[\mathit{plen}\ \mathfrak{J}F(x_o)\right]^T = \mathit{plen}\left[\mathfrak{J}F(x_o)\right]^T$.

2°) Montrons que $\partial F_v(x_o) = \left[\mathit{plen}\ \mathfrak{J}F(x_o)\right]^T \cdot v$.

Comme $\left[\mathit{plen}\ \mathfrak{J}F(x_o)\right]^T = \mathit{plen}\left[\mathfrak{J}F(x_o)\right]^T$, cette égalité est immédiate car

$\partial F_v(x_o) = \left[\mathfrak{J}F(x_o)\right]^T \cdot v$ et $\left[\mathfrak{J}F(x_o)\right]^T \cdot v = \left[\mathit{plen}\left[\mathfrak{J}F(x_o)\right]^T\right] \cdot v$.

3°) Montrons que $\left[\mathit{plen}\ \mathfrak{J}F(x_o)\right]^T$ est le plus grand ensemble convexe compact $\mathcal{K}(F;x_o)$ de $\mathcal{M}(q,p)$ vérifiant $\partial F_v(x_o) = \mathcal{K}(F;x_o)\cdot v$.

Soit $\mathcal{C}$ un ensemble convexe compact de $\mathcal{M}(q,p)$ tel que $\partial F_v(x_o) = \mathcal{C}^T \cdot v$. Il faut prouver que $\mathcal{C}$ est inclus dans $\mathit{plen}\ \mathfrak{J}F(x_o)$.

Par définition, la fonction de support de $\mathcal{C}^T \cdot v$ est la fonction s_v définie sur $\mathbb{R}^p$ et à valeurs réelles par

$$s_v(u) = \max\ \{\ \langle u, C^T \cdot v\rangle \mid C \in \mathcal{C}\ \}$$

pour tout u dans $\mathbb{R}^p$.

Il s'agit bien d'un maximum car $\mathcal{C}$ est un ensemble compact et le produit scalaire est une application continue.

Comme la fonction de support de $\partial F_v(x_o)$ est $F^\circ(x_o;\cdot,v)$, l'égalité $\partial F_v(x_o) = \mathcal{C}^T \cdot v$ entraîne que

$$s_v(u) = F^\circ(x_o;u,v)$$

pour tout point u de $\mathbb{R}^p$.

Autrement dit,

$$\max\ \{\ \langle u, C^T \cdot v\rangle \mid C \in \mathcal{C}\ \} = F^\circ(x_o;u,v)$$

pour tout u appartenant à $\mathbb{R}^p$.

Donc, comme v a été pris arbitrairement,

$$\langle u, C^T v\rangle \le F^{\circ}(x_o; u, v)$$

pour tout u dans $\mathbb{R}^p$ et pour toute matrice C de $\mathscr{C}$.

C'est-à-dire que

$$\langle Cu, v\rangle \le F^{\circ}(x_o; u, v).$$

pour tout u dans $\mathbb{R}^p$ et pour toute matrice C de $\mathscr{C}$.

Pour tout C appartenant à $\mathscr{C}$, le théorème 3.4.9 indique alors que C appartient à *plen* $\mathfrak{J}F(x_o)$. Donc $\mathscr{C}$ est inclus dans *plen* $\mathfrak{J}F(x_o)$.

□

Par cette propriété, nous achevons de voir que *plen* $\mathfrak{J}F(x_o)$ à les mêmes caractéristiques que $\mathfrak{J}F(x_o)$ sans en avoir les désavantages.

Pour compléter ce chapitre, nous allons décrire quelques propriétés mettant en relations les ensembles $\mathfrak{J}F(x_o)$, *plen* $\mathfrak{J}F(x_o)$ et $\partial f_i(x_o)$, où i est compris entre 1 et q.

Notons par $\left[\partial f_1(x_o) \ \ldots \ \partial f_q(x_o)\right]^T$ l'ensemble des matrices de $\mathcal{M}(q,p)$ dont la $i^{ème}$ ligne est formée d'un élément de $\partial f_i(x_o)$.

Propriété 3.4.12 : $\mathfrak{J}F(x_o) \subseteq$ *plen* $\mathfrak{J}F(x_o) \subseteq \left[\partial f_1(x_o) \ \ldots \ \partial f_q(x_o)\right]^T$.

Preuve :

Par définition de l'enveloppe pleine d'un ensemble, la première inclusion est vraie. Il reste donc à montrer la seconde, c'est-à-dire que *plen* $\mathfrak{J}F(x_o)$ est inclus dans $\left[\partial f_1(x_o) \;\ldots\; \partial f_q(x_o)\right]^T$.

1°) Commençons par voir que $\left[\partial f_1(x_o) \;\ldots\; \partial f_q(x_o)\right]^T$ est un ensemble convexe, compact et plein.

1.1°) Montrons que $\left[\partial f_1(x_o) \;\ldots\; \partial f_q(x_o)\right]^T$ est un ensemble convexe.

Considérons M et N appartenant à $\left[\partial f_1(x_o) \;\ldots\; \partial f_q(x_o)\right]^T$, c'est-à-dire que $M = [m_1 \ldots m_q]^T$ et $N = [n_1 \ldots n_q]^T$, où m_i et n_i appartiennent à $\partial f_i(x_o)$ pour tout i compris entre 1 et q.

Soit i tel que $1 \leq i \leq q$. Par la propriété 2.2.2 concernant le gradient généralisé, $\partial f_i(x_o)$ est un ensemble convexe. Ainsi, $(\lambda \cdot m_i + (1-\lambda) \cdot n_i)$ appartient à $\partial f_i(x_o)$ quel que soit λ appartenant à $[0,1]$.

Ceci étant vrai pour tout i compris entre 1 et q, il suit que $(\lambda \cdot M + (1-\lambda) \cdot N)$ appartient à $\left[\partial f_1(x_o) \;\ldots\; \partial f_q(x_o)\right]^T$ pour tout λ appartenant à $[0,1]$.

Donc $\left[\partial f_1(x_o) \;\ldots\; \partial f_q(x_o)\right]^T$ est un ensemble convexe.

1.2°) Montrons que $\left[\partial f_1(x_o) \;\ldots\; \partial f_q(x_o)\right]^T$ est un ensemble compact.

La propriété 2.2.2 concernant la gradient généralisé indique que $\partial f_i(x_o)$ est un ensemble compact quel que soit i compris entre 1 et q.

De plus, par le théorème 1.2.4 de Tichonoff, le produit cartésien d'espaces compacts est compact.

Donc, $\left[\partial f_1(x_0)\ \ldots\ \partial f_q(x_0)\right]^T$ est un ensemble compact.

1.3°) Montrons que $\left[\partial f_1(x_0)\ \ldots\ \partial f_q(x_0)\right]$ est un ensemble plein.

Soit M une matrice de $\mathcal{M}(q,p)$ telle que $M\cdot u$ appartient à $\left[\partial f_1(x_0)\ \ldots\ \partial f_q(x_0)\right]^T\cdot u$ quel que soit u dans $\mathbb{R}^p$. Il faut obtenir que M appartient à $\left[\partial f_1(x_0)\ \ldots\ \partial f_q(x_0)\right]^T$.

Posons $M = [m_1 \ldots m_q]^T$ et soit i compris entre 1 et q.

Comme $M\cdot u$ appartient à $\left[\partial f_1(x_0)\ \ldots\ \partial f_q(x_0)\right]^T\cdot u$ pour tout u dans $\mathbb{R}^p$, il suit que $m_i\cdot u$ appartient à $[\partial f_i(x_0)]^T\cdot u$ pour tout point u de $\mathbb{R}^p$. Or, par la propriéré 3.4.6, $\partial f_i(x_0)$ est un ensemble plein. D'où, m_i appartient à $\partial f_i(x_0)$.

Ceci étant vrai pour tout i compris entre 1 et q, nous obtenons que M est une matrice de $\left[\partial f_1(x_0)\ \ldots\ \partial f_q(x_0)\right]^T$.

Donc, $\left[\partial f_1(x_0)\ \ldots\ \partial f_q(x_0)\right]^T$ est un ensemble plein.

2°) Montrons que *plen* $\mathfrak{J}F(x_0)$ est inclus dans $\left[\partial f_1(x_0)\ \ldots\ \partial f_q(x_0)\right]^T$.

Soit M une matrice de *plen* $\mathfrak{J}F(x_0)$. Si $\left(plen\ \mathfrak{J}F(x_0)\right)\cdot u$ est inclus dans $\left[\partial f_1(x_0)\ \ldots\partial f_q(x_0)\right]^T\cdot u$ quel soit u dans $\mathbb{R}^p$, alors M appartient à $\left[\partial f_1(x_0)\ \ldots\ \partial f_q(x_0)\right]^T$ car $\left[\partial f_1(x_0)\ \ldots\ \partial f_q(x_0)\right]^T$ est un ensemble plein.

En appliquant la propriété 3.4.7, il suffit de montrer que

$$\mathfrak{J}F(x_0)\cdot u \text{ est contenu dans } \left[\partial f_1(x_0)\ \ldots\partial f_q(x_0)\right]^T\cdot u$$

pour tout point u de $\mathbb{R}^p$.

Soit u appartenant à $\mathbb{R}^p$. Par la définition 3.1.1, nous savons que

$$\mathfrak{J}F(x_o)\cdot u = \text{co}\ \{\ M\cdot u \in \mathbb{R}^q \mid M \in \mathfrak{J}(F;x_o)\ \}.$$

2.1°) Commençons par voir que

$$\{\ M\cdot u \in \mathbb{R}^q \mid M \in \mathfrak{J}(F;x_o)\ \} \subseteq \left[\partial f_1(x_o)\ \ldots\partial f_q(x_o)\right]^T\cdot u.$$

Soit M une matrice de $\mathfrak{J}(F;x_o)$ et posons $M = [m_1\ldots m_q]^T$. Nous devons montrer que $m_i\cdot u$ appartient à $[\partial f_i(x_o)]^T\cdot u$ pour tout i compris entre 1 et q.

Soit i compris entre 1 et q.

Comme M est une matrice de $\mathfrak{J}(F;x_o)$, il existe une suite $(x_n)_{n=1}^{\infty}$ de points de qui converge vers x_o telle que $M = \lim_{n\to\infty} JF(x_n)$.

Ainsi, $m_i = \lim_{n\to\infty} [\nabla f_i(x_n)]^T$. D'où, m_i appartient à $[\partial f_i(x_o)]^T$.

Donc, $m_i\cdot u$ appartient à $[\partial f_i(x_o)]^T\cdot u$.

2.2°) Montrons que $\mathfrak{J}F(x_o)\cdot u \subseteq \left[\partial f_1(x_o)\ \ldots\partial f_q(x_o)\right]^T\cdot u$.

Nous savons que $\left[\partial f_1(x_o)\ \ldots\partial f_q(x_o)\right]^T$ est un ensemble convexe, d'où $\left[\partial f_1(x_o)\ \ldots\ \partial f_q(x_o)\right]^T\cdot u$ est également un ensemble convexe.

De plus, nous venons de voir que

$$\{\ M\cdot u \in \mathbb{R}^q \mid M \in \mathfrak{J}(F;x_o)\ \} \subseteq \left[\partial f_1(x_o)\ \ldots\partial f_q(x_o)\right]^T\cdot u.$$

Or, par définition de l'enveloppe convexe d'un ensemble, $\mathfrak{J}F(x_o)\cdot u$ est le plus petit ensemble convexe contenant $\{\ M\cdot u \in \mathbb{R}^q \mid M \in \mathfrak{J}(F;x_o)\ \}$.

Donc, $\mathfrak{J}F(x_o)\cdot u$ est inclus dans $\left[\partial f_1(x_o)\ \ldots\partial f_q(x_o)\right]^T\cdot u$.

□

Dans la propriété qui suit, nous allons utilisé le symbole de Kronecker δ_{ij}, c'est-à-dire le symbole défini par

$$\delta_{ij} = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases} .$$

Propriété 3.4.13 : Soit e_i un point de $\mathbb{R}^q$ défini pour i compris entre 1 et q par $e_i = (\delta_{ij})_{j=1}^{q}$,

alors $\{ M\cdot e_i \in \mathbb{R}^p \mid M^T \in \mathfrak{J}F(x_o) \} = \partial f_i(x_o)$.

Preuve :

Soit la fonction p_i définie, sur $\mathbb{R}^q$ et à valeurs réelles ,par $p_i(y) = y_i$ pour tout $y = (y_i)_{i=1}^{q}$ appartenant à $\mathbb{R}^q$. Cette fonction est continûment différentiable et $\nabla p_i(y) = \nabla y_i = e_i$.

Par la propriété annexe A.4, nous savons que

$$\partial(p_i \circ F)(x_o) = [\mathfrak{J}F(x_o)]^T \cdot \nabla p_i(F(x_o)).$$

Or, $(p_i \circ F) = f_i$, d'où

$$\partial f_i(x_o) = [\mathfrak{J}F(x_o)]^T \cdot \nabla p_i(F(x_o)).$$

C'est-à-dire que

$$\partial f_i(x_o) = [\mathfrak{J}F(x_o)]^T \cdot e_i.$$

Ainsi, $\partial f_i(x_o) = \{ M\cdot e_i \in \mathbb{R}^p \mid M^T \in \mathfrak{J}F(x_o) \}$.

□

Cette propriété termine le chapitre à propos de la matrice Jacobienne généralisée. Principalement, nous retiendrons la définition de $\mathfrak{J}F(x_o)$ et ses caractéristiques, ainsi que la notion de bifonction de support de $\mathfrak{J}F(x_o)$.

Chapitre 4

MATRICE HESSIENNE GENERALISEE

A présent, nous allons appliquer la définition de la matrice Jacobienne généralisée vue au chapitre précédent à une certaine classe de fonctions. De cette manière, nous allons définir la matrice Hessienne généralisée pour cette classe de fonctions.

En plus des propriétés qui sont déduites du chapitre précédent, nous décrirons des extensions de propriétés qui existent chez les fonctions deux fois continûment différentiables.

Dans ce but, considérons une fonction f définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles.

4.1 Définition et premières propriétés.

Commençons par définir l'ensemble des fonctions qui seront employées.

Définition 4.1.1 : $C^{1,1}(\mathcal{O})$ est l'ensemble des fonctions f définies sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles telles que f est continûment différentiable sur $\mathcal{O}$ et ∇f est localement Lipschitzienne sur $\mathcal{O}$.

Pour une fonction f appartenant à $C^{1,1}(\mathcal{O})$, nous dirons également que f est de classe $C^{1,1}$ sur $\mathcal{O}$.

Cette classe de fonctions ne rend pas compte de toutes les fonctions qui ne sont pas deux fois continûment différentiables, mais elle permet d'élargir l'ensemble des fonctions deux fois continûment différentiables.

Enonçons quelques propriétés élémentaires de cette classe de fonctions.

Propriété 4.1.2 : Soient f_1 et f_2 des fonctions définies sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles.
Supposons que f_1 et f_2 appartiennent à $C^{1,1}(\mathcal{O})$, alors $(f_1 + f_2)$ appartient à $C^{1,1}(\mathcal{O})$.

Preuve :

Comme f_1 et f_2 sont des fonctions continûment différentiables sur $\mathcal{O}$, nous savons que $(f_1 + f_2)$ est une fonction continûment différentiable sur $\mathcal{O}$.

Comme ∇f_1 et ∇f_2 sont des fonctions localement Lipschitziennes sur $\mathcal{O}$, la propriété 1.6.2 indique que $\nabla(f_1 + f_2)$ est une fonction localement Lipschitzienne sur $\mathcal{O}$.

Donc $(f_1 + f_2)$ est de classe $C^{1,1}$ sur $\mathcal{O}$.

□

Propriété 4.1.3 : Soit f une fonction définie sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles et soit r un nombre réel. Supposons que f appartient à $C^{1,1}(\mathcal{O})$, alors $(r\cdot f)$ appartient à $C^{1,1}(\mathcal{O})$.

Preuve :

Comme f est une fonction continûment différentiable sur $\mathcal{O}$, nous savons que $(r\cdot f)$ est une fonction continûment différentiable sur $\mathcal{O}$. Comme ∇f est une fonction localement Lipschitzienne sur $\mathcal{O}$, la propriété 1.6.3 indique que $\nabla(r\cdot f)$ est une fonction localement Lipschitzienne sur $\mathcal{O}$.

Donc $(r\cdot f)$ appartient à $C^{1,1}(\mathcal{O})$.

□

Remarquons que, pour de telles fonctions, le gradient existe en tout point de $\mathcal{O}$ car ces fonctions sont continûment différentiables sur $\mathcal{O}$. De plus, nous pouvons alors appliquer la définiton 3.1.1 au gradient pour obtenir la matrice Hessienne généralisée car ces fonctions sont localement Lipschitzienne sur $\mathcal{O}$.

Avant de passer à la définition de la matrice Hessienne généralisée, notons par $\Omega^c_{\nabla f}$ le complémentaire dans $\mathcal{O}$ de l'ensemble $\Omega_{\nabla f}$ des points où ∇f n'est pas différentiable.

Définition 4.1.4: Soient f appartenant à $C^{1,1}(\mathcal{O})$ et x_o dans $\mathcal{O}$.
La <u>matrice Hessienne généralisée</u> de f au point x_o est l'ensemble, noté par $\partial^2 f(x_o)$, défini par l'enveloppe convexe des matrices M de $\mathcal{M}(p,p)$ telles que $M = \lim_{n\to\infty} \nabla^2 f(x_n)$ où $(x_n)_{n=1}^{\infty}$ est une suite de points de $\Omega^c_{\nabla f}$ qui converge vers x_o. C'est-à-dire que

$$\partial^2 f(x_o) = \text{co}\ [\partial^2(f,x_o)], \tag{4.1}$$

où $\partial^2(f,x_o) = \{\ M \in \mathcal{M}(p,p)\ \mid\ \exists (x_n)_{n=1}^{\infty} \subseteq \Omega^c_{\nabla f}$ convergente vers x_o telle que $M = \lim_{n\to\infty} \nabla^2 f(x_n)\ \}$.

Comme $\partial^2 f(x_o)$ est une particularisation de la matrice Jacobienne généralisée développée au chapitre précédent, nous avons directement les propriétés suivantes.

Considérons x_o appartenant à $\mathcal{O}$.

Propriété 4.1.5 : $\partial^2 f(x_o)$ est un ensemble non vide, compact et convexe.

Propriété 4.1.6 : L'application

$$\partial^2 f:\ \mathcal{O} \subseteq \mathbb{R}^p \longrightarrow \mathcal{P}(\mathcal{M}(p,p))$$
$$x \longrightarrow \partial^2 f(x)$$

est localement bornée.

Propriété 4.1.7 : L'application $\partial^2 f : \mathcal{O} \subseteq \mathbb{R}^p \longrightarrow \mathcal{P}(\mathcal{M}(p,p))$

$$x \longrightarrow \partial^2 f(x)$$

est semi-continue supérieurement.

Propriété 4.1.8 : Si ∇f est strictement différentiable en x_o, alors $\partial^2 f(x_o) = \{ D^2 f(x_o) \}$, où $D^2 f(x_o)$ désigne la différentielle stricte de ∇f au point x_o.

Maintenant, donnons deux exemples concernant la matrice Hessienne généralisée.

Exemple 4.1.9 :

Soit g une fonction deux fois continûement différentiable définie sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles.
Considérons la fonction f définie sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles par

$$f(x) = [g^+(x)]^2,$$

pour tout x dans $\mathcal{O}$, avec $g^+(x) = \max \{ g(x), 0 \}$.
Alors, nous calculons que

$$\partial^2 f(x_o) = \begin{cases} \{ 0 \} & \text{si } g(x_o) < 0 \\ \{ 2\cdot\nabla^2 g(x_o)\cdot g(x_o) + 2\cdot\nabla g(x_o)\cdot[\nabla g(x_o)]^T \} & \text{si } g(x_o) > 0. \\ \{ 2\cdot\alpha\cdot\nabla g(x_o)\cdot[\nabla g(x_o)]^T \mid \alpha \in [0,1] \} & \text{si } g(x_o) = 0 \end{cases}$$

Preuve :
Avant de pouvoir considérer $\partial^2 f(x_o)$, il faut voir que f est de classe $C^{1,1}(\mathcal{O})$.

1°) f est de classe $C^{1,1}(\mathcal{O})$.

Réécrivons la définition de la fonction f.

$$f(x) = \begin{cases} 0 & \text{si } g(x) < 0 \\ [g(x)]^2 & \text{si } g(x) \geq 0 \end{cases}.$$

Ainsi, si $g(x) < 0$, alors f est deux fois continûment différentiable en x car $f(x) = 0$.

De même, si $g(x) > 0$, alors f est deux fois continûment différentiable en x car $f(x) = [g(x)]^2$, et le composé de fonctions deux fois continûment différentiables est deux fois continûment différentiable.

D'où, les seuls points où le caractère $C^{1,1}$ doit être vérifié sont les points x_o tels que $g(x_o) = 0$.

1.1°) f est différentiable en x_o et $\nabla f(x_o) = 0$.

Il faut voir que

$$\lim_{x \to x_o} \frac{f(x) - f(x_o) - [\nabla f(x_o)]^T \cdot (x - x_o)}{\|x - x_o\|} = 0.$$

Par définition de f, nous savons que $f(x_o) = 0$. De plus, nous devons montrer que $\nabla f(x_o) = 0$, d'où, il suffit de prouver que

$$\lim_{x \to x_o} \frac{f(x)}{\|x - x_o\|} = 0.$$

Autrement dit, par le critère de réduction aux suites, il faut obtenir que, pour toute suite $(x_n)_{n=1}^{\infty}$ de points de $\mathcal{O}$ qui converge vers x_o,

$$\lim_{n \to \infty} \frac{f(x_n)}{\|x_n - x_o\|} = 0.$$

Soit $(x_n)_{n=1}^{\infty}$ une telle suite.

Remarquons que, pour tout x dans $\mathcal{O}$, nous avons

$$0 \le f(x) \le [g(x)]^2.$$

D'où,

$$0 \le \frac{f(x_n)}{\|x_n - x_o\|} \le \frac{[g(x_n)]^2}{\|x_n - x_o\|}.$$

Or, $[g(\cdot)]^2$ est une fonction deux fois continûment différentiable car elle est le résultat de la composition de fonctions deux fois continûment différentiables, et nous calculons, grâce aux règles de différentiation, que $\nabla[g(x)]^2 = 2 \cdot g(x) \cdot \nabla g(x)$ quel que soit x dans $\mathcal{O}$.

En x_o, nous avons donc que $\nabla[g(x_o)]^2 = 0$.

D'où,

$$\lim_{n\to\infty} \frac{[g(x_n)]}{\|x_n - x_o\|} = 0,$$

donc

$$\lim_{n\to\infty} \frac{f(x_n)}{\|x_n - x_o\|} = 0.$$

D'une façon générale, nous pouvons exprimer maintenant la forme de la fonction ∇f. En effet, pour x tel que $g(x) = 0$, nous connaissons son expression. Et pour x tel que $g(x)$ est différent de 0, nous pouvons calculer $\nabla f(x)$ sans difficulté. Ainsi, nous obtenons

$$\nabla f(x) = \begin{cases} 0 & \text{si } g(x) < 0 \\ 2 \cdot \nabla g(x) \cdot g(x) & \text{si } g(x) \ge 0 \end{cases}.$$

1.2°) ∇f est une fonction localement Lipschitzienne en x_o.

Il faut trouver $\varepsilon > 0$ et $K > 0$ tels que

$$\| \nabla f(x) - \nabla f(y) \| \leq K \cdot \| x - y \|$$

pour tout x et y appartenant à $B(x_o, \varepsilon)$.

Pour que cette inégalité soit vérifiée, trois cas sont à envisager:

i) $g(x) \leq 0$ et $g(y) \leq 0$,

ii) $g(x) > 0$ et $g(y) \leq 0$,

iii) $g(x) > 0$ et $g(y) > 0$.

i) Supposons que les points x et y envisagés soient tels que $g(x) \leq 0$ et $g(y) \leq 0$.

Nous avons alors que $\nabla f(x) = 0$ et $\nabla f(y) = 0$. D'où,

$$\| \nabla f(x) - \nabla f(y) \| = 0.$$

Dans ce cas, nous pouvons prendre n'importe quelles constantes $\varepsilon > 0$ et $K > 0$.

En effet, nous avons que

$$\| \nabla f(x) - \nabla f(y) \| \leq K \cdot \| x - y \|,$$

quels que soient x et y appartenant à $B(x_o, \varepsilon)$ tels que $g(x) \leq 0$ et $g(y) \leq 0$.

Notes: Avant de développer les deux derniers cas, énonçons clairement les hypothèses dont nous aurons besoin.

g est une fonction deux fois continûment différentiable en x_o, d'où,

- les fonctions g et ∇g sont localement Lipschitziennes en x_o.

Autrement dit,

\# il existe des constantes $\varepsilon_g > 0$ et $K_g > 0$ telles que

$\| g(x) - g(y) \| \le K_g \cdot \| x - y \|$

pour tout x et y appartenant à $B(x_o, \varepsilon_g)$, et

\# il existe des constantes $\varepsilon_{\nabla g} > 0$ et $K_{\nabla g} > 0$ telles que

$\| \nabla g(x) - \nabla g(y) \| \le K_{\nabla g} \cdot \| x - y \|$

pour tout x et y appartenant à $B(x_o, \varepsilon_{\nabla g})$.

■ les fonctions g et ∇g sont continues en x_o. Ainsi, ces fonctions sont bornées sur des boules fermées contenues dans l'ouvert $\mathcal{O}$. Autrement dit,

\# il existe des constantes $\eta_g > 0$ et $C_g > 0$ telles que

$\| g(x) \| \le C_g$

pour tout x dans $B(x_o, \eta_g)$, et

\# il existe des constantes $\eta_{\nabla g} > 0$ et $C_{\nabla g} > 0$ telles que

$\| g(x) \| \le C_{\nabla g}$

pour tout x dans $B(x_o, \eta_{\nabla g})$.

A présent, envisageons les points suivants.

ii) Supposons que les points x et y envisagés soient tels que $g(x) > 0$ et $g(y) \le 0$.

Ainsi, $\nabla f(x) = 2 \cdot \nabla g(x) \cdot g(x)$ et $\nabla f(y) = 0$. D'où,

$\| \nabla f(x) - \nabla f(y) \| = 2 \cdot \| \nabla g(x) \cdot g(x) \|$.

Comme la fonction g est continue, le théorème de la valeur intermédiaire indique qu'il existe un nombre λ appartenant à l'intervalle $[0,1]$ pour lequel le point $(\lambda \cdot x + (1-\lambda) \cdot y)$, noté par z, est tel que $g(z) = 0$.

Ainsi,

$$\| \nabla f(x) - \nabla f(y) \| = 2\cdot\| \nabla g(x)\cdot g(x) - \nabla g(z)\cdot g(z) \|.$$

L'inégalité triangulaire implique alors que

$$\| \nabla f(x) - \nabla f(y) \| \le 2\cdot\| \nabla g(x)\cdot g(x) - \nabla g(z)\cdot g(x) \| + 2\cdot\| \nabla g(z)\cdot g(x) - \nabla g(z)\cdot g(z) \|.$$

D'où, compte tenu de l'inégalité de Cauchy-Schwarz,

$$\| \nabla f(x) - \nabla f(y) \| \le 2\cdot\| \nabla g(x) - \nabla g(z) \|\cdot\| g(x) \| + 2\cdot\| \nabla g(z) \|\cdot\| g(x) - g(z) \|.$$

Choisissons $\varepsilon = \min\ (\varepsilon_g, \varepsilon_{\nabla g}, \eta_g, \eta_{\nabla g})$. Ainsi, si x et y appartiennent à $B(x_o,\varepsilon)$, alors

$$\| \nabla f(x) - \nabla f(y) \| \le 2\cdot(K_{\nabla g}\cdot C_g + C_{\nabla g}\cdot K_g)\cdot\| x - z \|.$$

Or, $\| x - z \| \le \| x - y \|$ car $\| x - z \| = (1-\lambda)\cdot\| x - y \|$ et λ est un nombre de l'intervalle réel $[0,1]$. Donc,

$$\| \nabla f(x) - \nabla f(y) \| \le 2\cdot(K_{\nabla g}\cdot C_g + C_{\nabla g}\cdot K_g)\cdot\| x - y \|.$$

Finalement, en choisissant $K = 2\cdot(K_{\nabla g}\cdot C_g + C_{\nabla g}\cdot K_g)$, nous obtenons que, si $g(x) > 0$ et $g(y) \le 0$, la fonction ∇f est localement Lipschitzienne x_o.

iii) Supposons que les points x et y envisagés soient tels que $g(x) > 0$ et $g(y) > 0$.

Ainsi, $\nabla f(x) = 2\cdot\nabla g(x)\cdot g(x)$ et $\nabla f(y) = 2\cdot\nabla g(y)\cdot g(y)$. D'où,

$$\| \nabla f(x) - \nabla f(y) \| = 2\cdot\| \nabla g(x)\cdot g(x) - \nabla g(y)\cdot g(y) \|.$$

Par le même mécanisme que le point précédent, en choisissant $\varepsilon = \min\ (\varepsilon_g, \varepsilon_{\nabla g}, \eta_g, \eta_{\nabla g})$ et $K = 2\cdot(K_{\nabla g}\cdot C_g + C_{\nabla g}\cdot K_g)$, nous obtenons que

$$\| \nabla f(x) - \nabla f(y) \| \le 2\cdot(K_{\nabla g}\cdot C_g + C_{\nabla g}\cdot K_g)\cdot\| x - y \|,$$

pour tout x et y dans $B(x_o,\varepsilon)$ et tels que $g(x) > 0$ et $g(y) > 0$.

C'est-à-dire que, dans ce cas, la fonction ∇f est localement Lipschitzienne en x_o.

Donc, en choisissant les constantes $\varepsilon = \min\ (\varepsilon_g, \varepsilon_{\nabla g}, \eta_g, \eta_{\nabla g})$ et $K = 2 \cdot (K_{\nabla g} \cdot C_g + C_{\nabla g} \cdot K_g)$, nous obtenons que la fonction ∇f est localement Lipschitzienne en x_o.

En <u>conclusion</u>, f est une fonction différentiable sur $\mathcal{O}$ et ∇f est une fonction localement Lipischitzienne sur $\mathcal{O}$, c'est-à-dire f est de classe $C^{1,1}$ sur $\mathcal{O}$.

2°) Calculons l'expression de $\partial^2 f(x_o)$.

Nous allons diviser ce calcul en trois parties:

i) $g(x_o) > 0$,

ii) $g(x_o) < 0$,

iii) $g(x_o) = 0$.

i) Supposons que $g(x_o) > 0$.

Ainsi, $f(x_o) = [g(x_o)]^2$. D'où, f est deux fois continûment différentiable en x_o car f est alors composé de fonctions deux fois continûment différentiables en x_o. Les règles de différentiabilité entraînent alors que

$$\nabla^2 f(x_o) = 2 \cdot \nabla^2 g(x_o) \cdot g(x_o) + 2 \cdot \nabla g(x_o) \cdot [\nabla g(x_o)]^T.$$

Compte tenu de la propriété 2.2.3, nous obtenons que

$$\partial^2 f(x_o) = \{\ 2 \cdot \nabla^2 g(x_o) \cdot g(x_o) + 2 \cdot \nabla g(x_o) \cdot [\nabla g(x_o)]^T\ \}.$$

ii) Supposons que $g(x_o) < 0$.

Ainsi, $f(x_o) = 0$. D'où, f est deux fois continûment différentiable en x_o, et nous avons que

$$\nabla^2 f(x_o) = 0 \text{ (dans } \mathcal{M}(p,p)).$$

Compte tenu de la propriété 2.2.3, nous obtenons que

$$\partial^2 f(x_o) = \{ 0 \} \text{ (dans } \mathcal{M}(p,p)).$$

iii) Supposons que $g(x_o) = 0$.

Ainsi, f n'est plus nécessairement deux fois continûment différentiable en x_o. Nous savons uniquement que f est une fonction de classe $C^{1,1}$.

Par définition de $\partial^2 f(x_o)$, nous devons trouver l'enveloppe convexe des matrices M de $\mathcal{M}(p,p)$ telles que $M = \lim_{n\to\infty} \nabla^2 f(x_n)$, où $(x_n)_{n=1}^{\infty}$ est une suite de points de $\Omega^c_{\nabla f}$ qui converge vers x_o.

Or, les deux points précédents indiquent que

$$\Omega^c_{\nabla f} = \{ x \in \mathbb{R}^p \mid g(x) \neq 0 \}.$$

Ainsi, soit $\nabla^2 f(x_n) = \nabla^2 g(x_n)\cdot g(x_n) + 2\cdot\nabla g(x_n)\cdot[\nabla g(x_n)]^T$,

soit $\nabla^2 f(x_n) = 0$ (dans $\mathcal{M}(p,p)$).

Par passage à la limite, nous obtenons que

soit $M = 2\cdot\nabla g(x_o)\cdot[\nabla g(x_o)]^T$,

soit $M = 0$ (dans $\mathcal{M}(p,p)$).

Donc, $\partial^2 f(x_o) = \{ 2\cdot\alpha\cdot\nabla g(x_o)\cdot[\nabla g(x_o)]^T \mid \alpha \in [0,1] \}$.

En <u>conclusion</u>,

$$\partial^2 f(x_o) = \begin{cases} \{ 0 \} & \text{si } g(x_o) < 0 \\ \{ 2\cdot\nabla^2 g(x_o)\cdot g(x_o) + 2\cdot\nabla g(x_o)\cdot[\nabla g(x_o)]^T \} & \text{si } g(x_o) > 0. \\ \{ 2\cdot\alpha\cdot\nabla g(x_o)\cdot[\nabla g(x_o)]^T \mid \alpha \in [0,1] \} & \text{si } g(x_o) = 0 \end{cases}$$

■

Dans l'exemple suivant, nous montrons que les fonctions de classe $C^{1,1}$ apparaissent naturellement dans les problèmes d'optimisation.

Exemple 4.1.10 :

Soit m un nombre réel supérieur à un.

Considérons le problème d'optimisation suivant.

$$(P)\quad \begin{cases} \text{Minimiser } f_0(x) \\ \text{sous conditions : } f_1(x) \le 0, \\ \qquad\qquad\qquad\quad \vdots \\ \qquad\qquad\qquad\quad f_m(x) \le 0, \\ \qquad\qquad\qquad\quad x \in \mathbb{R}^p. \end{cases}$$

où, pour tout i compris entre 0 et m, les fonctions f_i sont définies sur $\mathbb{R}^p$ et à valeurs réelles, et sont deux fois continûment différentiables.

Soit r un nombre réel positif.

Dans [8], R.T. Rockaffellar décrit le Lagrangien augmenté L_r défini sur $\mathbb{R}^p \times \mathbb{R}^m$ et à valeurs réelles par

$$L_r(x,y) = f_0(x) + \sum_{i=1}^{m} \left\{ \left[[y_i + 2\cdot r\cdot f_i(x)]^+ \right]^2 - y_i^2 \right\},$$

pour tout x appartenant à $\mathbb{R}^p$ et $y = (y_i)_{i=1}^m$ appartenant à $\mathbb{R}^m$.

Montrons que L_r est de classe $C^{1,1}$ sur $\mathbb{R}^{p+m}$.

En effet, nous savons que les fonctions f_0^* et c_i, définies par $f_0^*(x,y) = f_0(x)$ et $c_i(x,y) = y_i^2$ pour tout x dans $\mathbb{R}^p$ et y dans $\mathbb{R}^m$, sont deux fois continûment différentiables sur $\mathbb{R}^{p+m}$. Ainsi, ces fonctions sont de classes $C^{1,1}$ sur $\mathbb{R}^{p+m}$.

De plus, les fonctions ϕ_i définies par $\phi_i(x,y) = y_i + 2\cdot r\cdot f_i(x)$ pour tout x appartenant à $\mathbb{R}^p$ et y appartenant à $\mathbb{R}^m$ sont deux fois continûment différentiables sur $\mathbb{R}^{p+m}$ car chaque f_i est deux fois continûment différentiable sur $\mathbb{R}^p$.

Ainsi, l'exemple 4.1.9 implique que $[\phi_i^{+}(\cdot,\cdot)]^2$ appartient à $C^{1,1}(\mathbb{R}^{p+m})$.

D'où, comme L_r est formé d'une somme de fonctions qui appartiennent à $C^{1,1}(\mathbb{R}^{p+m})$, la propriété 4.1.2 entraîne que L_r appartient également à $C^{1,1}(\mathbb{R}^{p+m})$.

Donc, nous pouvons envisager la matrice Hessienne généralisée du Lagrangien augmenté. Mais, remarquons que, en tout point (x,y) de $\mathbb{R}^p \times \mathbb{R}^m$ tel que $(y_i + 2\cdot r\cdot f_i(x))$ est non nul pour tout i compris entre 1 et m, l'exemple 4.1.9 nous indique que le Lagrangien augmenté est deux fois continûment différentiable. Nous n'enviagerons donc la matrice Hessienne généralisée que pour les points d'un hyperplan d'équation $(y_i + 2\cdot r\cdot f_i(x)) = 0$.

Pour de tels points, nous obtenons

$\partial^2_x L_r(x,y)$ la matrice Hessienne généralisée pour la fonction $L_r(\cdot,y)$ au point (x,y),

$\partial^2_y L_r(x,y)$ la matrice Hessienne généralisée pour la fonction $L_r(x,\cdot)$ au point (x,y), et

$\partial^2 L_r(x,y)$ la matrice Hessienne généralisée pour la fonction $L_r(\cdot,\cdot)$ au point (x,y).

Maintenant, envisageons le problème dual de *(P)* associé au Lagrangien augmenté L_r.

$$(Dr) \begin{cases} \text{Maximiser } g_r(y) \\ \text{sous condition } y \geq 0 \end{cases},$$

où $g_r(y) = \inf_{x\in\mathbb{R}^p} L_r(x,y)$.

Dans le cas où les fonctions f_i sont convexes quel que soit i et où $r > 0$, la propriété annexe D.1 entraîne que la fonction g_r est continûment différentiable sur $\mathbb{R}^m$,et que la fonction ∇g_r est Lipschitzienne de constante de Lipischitz $(\frac{1}{2 \cdot r})$.
Donc g_r appartient à $C^{1,1}(\mathbb{R}^m)$.
Ainsi, dans cet exemple, nous voyons que des fonctions de classe $C^{1,1}$ existent dans la pratique. Ces fonctions $C^{1,1}$ peuvent survenir assez rapidement. En effet, au départ de notre exemple, toutes les fonctions sont deux fois continûment différentiables. Cependant, le Lagrangien augmenté perd cette propriété pour n'être plus que de classe $C^{1,1}$. De même, la fonction objectif du problème dual n'est également que de classe $C^{1,1}$.
La propriété 4.1.8 impliquant que la matrice Hessienne généralisée est une bonne extension de la matrice Hessienne, il est intéressant de pouvoir déterminer les ensembles $\partial^2_x L_r(x,y)$, $\partial^2_y L_r(x,y)$ et $\partial^2 L_r(x,y)$.

■

4.2 Règles de calcul.

Soit x_o appartenant à $\mathcal{O}$ et supposons que f soit de classe $C^{1,1}$ sur $\mathcal{O}$.

Considérons la bifonction de support de $\partial^2 f(x_o)$. Par le chapitre précédent, il s'agit de la fonction $\nabla f^\circ(x_o;\cdot,\cdot)$ définie par $\nabla f^\circ(x_o;u,v) = \max \{ \langle Mu,v \rangle \mid M \in \partial^2 f(x_o) \}$ quels que soient u et v dans $\mathbb{R}^p$.

Par la suite nous noterons cette bifonction de support de $\partial^2 f(x_o)$ par $f^{\circ\circ}(x_o;\cdot,\cdot)$, c'est-à-dire que, pour tout u et v de $\mathbb{R}^p$,

$f^{\circ\circ}(x_o;u,v) = \max \{ \langle Mu,v\rangle \mid M \in \partial^2 f(x_o) \}$.

Le théorème 3.2.4 appliqué à $\partial^2 f(x_o)$ conduit à la propriété suivante.

Propriété 4.2.1 : Pour tout u et v appartenant à $\mathbb{R}^p$,

$$f^{\circ\circ}(x_o;u,v) = \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle \nabla f(x+\lambda u) - \nabla f(x) , v \rangle}{\lambda}.$$

De plus, nous avons la propriété particulière de $\partial^2 f(x_o)$ qui suit.

Propriété 4.2.2 : Pour tout u et v appartenant à $\mathbb{R}^p$,

$$f^{\circ\circ}(x_o;u,v) = f^{\circ\circ}(x_o;v,u).$$

Preuve :

Commençons par remarquer que toutes les matrices M de $\partial^2 f(x_o)$ sont symétriques.

En effet, soit x appartenant à $\Omega^c_{\nabla f}$, nous savons que $\nabla^2 f(x)$ est une matrice symétrique de $\mathcal{M}(p,p)$.

Ainsi, M est limite de matrices symétriques. Donc, M est une matrice symétrique.

Soient u et v appartenant à $\mathbb{R}^p$ et soit M appartenant à $\partial^2 f(x_o)$.

Nous savons que $\langle Mu,v\rangle = \langle u,M^T v\rangle$. Or, $M = M^T$ car M est une matrice symétrique. D'où, $\langle Mu,v\rangle = \langle u,Mv\rangle$.

Donc, $\max \{ \langle Mu,v\rangle \mid M \in \partial^2 f(x_o) \} = \max \{ \langle u,Mv\rangle \mid M \in \partial^2 f(x_o) \}$.

Autrement dit, $f^{\circ\circ}(x_o;u,v) = f^{\circ\circ}(x_o;v,u)$.

□

Exemple 4.2.3 :

Reprenons la fonction f définie dans l'exemple 4.1.9, c'est-à-dire que, pour tout x appartenant à $\mathcal{O}$,

$$f(x) = \begin{cases} 0 & \text{si } g(x) \le 0 \\ [g(x)]^2 & \text{si } g(x) \ge 0 \end{cases},$$

où g est une fonction deux fois continûment différentiable définie sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles.

Nous savons que f est de classe $C^{1,1}$ sur $\mathcal{O}$. Calculons la bifonction de support de $\partial^2 f(x_o)$.

Soit x_o dans $\mathcal{O}$ et supposons que $g(x_o) = 0$ (pour les mêmes raisons qu'à l'exemple 4.1.9, les autres cas sont inintéressants).

Considérons u et v deux points de $\mathbb{R}^p$.

Calculons l'expression de $f^{\circ\circ}(x_o;u,v)$.

Par définition , nous avons

$$f^{\circ\circ}(x_o;u,v) = \max \{ \langle Mu,v\rangle \mid M \in \partial^2 f(x_o) \}.$$

Or, l'exemple 4.1.9 indique que

$$\partial^2 f(x_o) = \{ 2\cdot\alpha\cdot\nabla g(x_o)\cdot[\nabla g(x_o)]^T \mid \alpha \in [0,1] \}.$$

D'où,

$$f^{\circ\circ}(x_o;u,v) = \max \{ \langle 2\cdot\alpha\cdot\nabla g(x_o)\cdot[\nabla g(x_o)]^T\cdot u,v\rangle \mid \alpha \in [0,1] \}.$$

Remarquons que $0 \leq f^{\circ\circ}(x_o;u,v)$.

En effet, pour $\alpha = 0$, l'expression $\langle 2\cdot\alpha\cdot\nabla g(x_o)\cdot[\nabla g(x_o)]^T\cdot u,v\rangle$ est nulle. D'où, en prenant le maximum de cette expression pour α dans l'intervalle réel $[0,1]$, nous obtenons nécessairement un nombre positif.

Mais rien ne nous permet d'affirmer que $\langle\nabla g(x_o)\cdot[\nabla g(x_o)]^T\cdot u,v\rangle$ est un nombre positif. Ainsi, pour pouvoir mettre cette constante en évidence du maximum, nous devons prendre le maximum entre la constante $\langle\nabla g(x_o)\cdot[\nabla g(x_o)]^T\cdot u,v\rangle$ et 0. Autrement dit, c'est $[\langle\nabla g(x_o)\cdot[\nabla g(x_o)]^T\cdot u,v\rangle]^+$ que nous devons mettre en évidence.

D'où,

$$f^{\circ\circ}(x_o;u,v) = 2\cdot[\langle\nabla g(x_o)\cdot[\nabla g(x_o)]^T\cdot u,v\rangle]^+\cdot \max\{\alpha \mid \alpha \in [0,1]\}.$$

C'est-à-dire que

$$f^{\circ\circ}(x_o;u,v) = 2\cdot[\langle\nabla g(x_o)\cdot[\nabla g(x_o)]^T\cdot u,v\rangle]^+.$$

En réécrivant ce résultat, nous obtenons finalement que

$$f^{\circ\circ}(x_o;u,v) = 2\cdot[\langle\nabla g(x_o),u\rangle\cdot\langle\nabla g(x_o),v\rangle]^+.$$

■

A présent rappelons nous que nous pouvons caractériser en partie $\partial^2 f(x_o)$ par sa bifonction de support. Mais, en tant que fonction de support de $\partial^2 f(x_o)\cdot u$, la propriété 3.3.3 indique que la bifonction de support de $\partial^2 f(x_o)$ détermine les éléments de $\partial^2 f(x_o)\cdot u$ quel que soit u dans $\mathbb{R}^p$.

Cette caractéristique va être utilisée pour évaluer la matrice Hessienne généralisée d'un composé de fonctions de classe $C^{1,1}$.

Commençons par poser les fonctions suivantes. Soit F définie sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}^q$, et soit ϕ définie sur l'ouvert $\mathcal{O}'$ de $\mathbb{R}^q$ et à valeurs réelles. Supposons que $F(\mathcal{O})$ soit inclus dans $\mathcal{O}'$, et notons par $(f_1,\ldots,f_q)^T$ les q composantes de F.

Dans les propriétés que nous allons développer, chaque composante f_i de F est supposée de classe $C^{1,1}$ sur $\mathcal{O}$, et la fonction ϕ est supposée de classe $C^{1,1}$ sur $\mathcal{O}'$. Ainsi, nous avons les remarques qui suivent.

Remarques :

(R1) Il existe des constantes $\varepsilon_1 > 0$ et $K_1 > 0$ telles que

$$\| \nabla\phi(F(x)) - \nabla\phi(F(y)) \| \le K_1 \cdot \| x - y \|$$

pour tout x et y appartenant à $B(x_o,\varepsilon_1)$.

Comme Φ appartient à $C^{1,1}(\mathcal{O}')$, la fonction $\nabla\phi$ est localement Lipschitzienne en $F(x_o)$. Autrement dit, il existe des constantes $\varepsilon_\phi > 0$ et $K_\phi > 0$ telles que

$$\| \nabla\phi(w) - \nabla\phi(z) \| \le K_\phi \cdot \| w - z \|$$

pour tout w et z appartenant à $B(F(x_o),\varepsilon_\phi)$.

Or, comme chaque composante de F appartient à $C^{1,1}(\mathcal{O})$, nous avons, d'une part que la fonction F est continue x_o. D'où, il existe une constante $\eta_F > 0$ telle que x appartient à $B(x_o,\eta_F)$ entraîne que F(x) appartient à $B(F(x_o),\varepsilon_\phi)$.

D'autre part, la fonction F est localement Lipschitzienne en x_o. Autrement dit, il existe des constantes $\gamma_F > 0$ et $L_F > 0$ telles que

$$\| F(x) - F(y) \| \le L_F \cdot \| x - y \|$$

pour tout x et y dans $B(x_o,\gamma_F)$.

D'où, en choisissant $\varepsilon_1 = \min(\eta_F, \gamma_F)$, nous obtenons que

$$\| \nabla\phi(F(x)) - \nabla\phi(F(y)) \| \le K_\phi \cdot L_F \cdot \| x - y \|$$

pour tout x et y appartenant à $B(x_o, \varepsilon_1)$.

Finalement, en prenant $K_1 = K_\phi \cdot L_F$, la remarque *(R1)* est vérifiée.

(R2) Il existe des constantes $\upsilon_F > 0$ et $C_F > 0$ telles que

$$\| JF(x) \| \le C_F$$

pour tout x dans $B(x_o, \upsilon_F)$.

Comme chaque composante f_i de F est de classe $C^{1,1}$ sur $\mathcal{O}$, la fonction JF est continue en x_o. Ainsi, il existe une boule fermée centrée en x_o sur laquelle JF est bornée.

(R3) Il existe des constantes $\upsilon_\phi > 0$ et $C_\phi > 0$ telles que

$$\| \nabla\phi(z) \| \le C_\phi$$

pour tout z dans $B(F(x_o), \upsilon_\phi)$.

Comme Φ est de classe $C^{1,1}$ sur $\mathcal{O}'$, la fonction $\nabla\phi$ est continue en $F(x_o)$. Ainsi, il existe une boule fermée centrée en $F(x_o)$ sur laquelle $\nabla\phi$ est bornée.

(R4) Il existe des constantes $\varepsilon_F > 0$ et $K_F > 0$ telles que

$$\| JF(x) - JF(y) \| \le K_F \cdot \| x - y \|$$

pour tout x et y appartenant à $B(x_o, \varepsilon_F)$.

Cette remarque est déduite du caractère $C^{1,1}$ de chaque composante f_i de F.

Ces remarques vont être utilisées dans les deux propriétés que nous allons décrire.

Propriété 4.2.4 : Sous les hypothèses énoncées ci-dessus qui concernent les fonctions F et ϕ, $(\phi \circ F)$ est de classe $\mathcal{C}^{1,1}$ sur $\mathcal{O}$.

Preuve :

1°) Remarquons que

$(\phi \circ F)$ est continûment différentiable sur $\mathcal{O}$ car le composé de fonctions continûment différentiables est continûment différentiable.

2°) Montrons que la fonction $\nabla(\phi \circ F)$ est localement Lipschitzienne en x_o.

Il faut trouver des constantes $\varepsilon > 0$ et $K > 0$ telles que

$$\| \nabla(\phi \circ F)(x) - \nabla(\phi \circ F)(y) \| \leq K \cdot \| x - y \|$$

pour tout x et y dans $B(x_o, \varepsilon)$.

Par le point précédent, nous savons que $\nabla(\phi \circ F)(z)$ existe quel que soit z dans $\mathcal{O}$, et nous calculons, grâce aux règles de différentiation, que $\nabla(\phi \circ F)(z) = [JF(z)]^T \cdot \nabla\phi(F(z))$.

Ainsi,

$$\| \nabla(\phi \circ F)(x) - \nabla(\phi \circ F)(y) \|$$
$$= \| [JF(x)]^T \cdot \nabla\phi(F(x)) - [JF(y)]^T \cdot \nabla\phi(F(y)) \| .$$

D'où, suite à l'inégalité triangulaire,

$$\| \nabla(\phi \circ F)(x) - \nabla(\phi \circ F)(y) \|$$
$$\leq \| [JF(x)]^T \cdot \nabla\phi(F(x)) - [JF(x)]^T \cdot \nabla\phi(F(y)) \|$$
$$+ \| [JF(x)]^T \cdot \nabla\phi(F(y)) - [JF(y)]^T \cdot \nabla\phi(F(y)) \| .$$

Et, par l'inégalité de Cauchy-Schwarz,

$$\| \nabla(\phi \circ F)(x) - \nabla(\phi \circ F)(y) \|$$
$$\leq \| [JF(x)] \| \cdot \|\nabla\phi(F(x)) - \nabla\phi(F(y)) \|$$
$$+ \| [JF(x)] - [JF(y)] \| \cdot \| \nabla\phi(F(y)) \| .$$

A chacun des ces facteurs, appliquons les remarques décrites avant cette propriété. Ainsi, en posant $\varepsilon = \min (\varepsilon_1, \upsilon_F, \upsilon_\phi, \varepsilon_F)$, nous obtenons

$$\| \nabla(\phi \circ F)(x) - \nabla(\phi \circ F)(y) \| \leq (C_F \cdot K_1 + K_F \cdot C_\phi) \cdot \| x - y \|$$

pour tout x et y dans $B(x_o, \varepsilon)$.

Donc, en, prenant $K = C_F \cdot K_1 + K_F \cdot C_\phi$, la fonction $\nabla(\phi \circ F)$ est localement Lipschitzienne en x_o.

Finalement, $(\phi \circ F)$ étant une fonction continûment différentiable sur $\mathcal{O}$ et $\nabla(\phi \circ F)$ étant une fonction localement Lipschitzienne sur $\mathcal{O}$, nous avons que $(\phi \circ F)$ est une fonction de classe $C^{1,1}$ sur $\mathcal{O}$.

□

Cette propriété nous permet de pouvoir envisager l'ensemble $\partial^2(\phi \circ F)(x_o)$. Ainsi, nous avons le théorème qui suit.

Théorème 4.2.5 : Pour tout u et v appartenant à $\mathbb{R}^p$,

$$(\phi \circ F)^{\circ\circ}(x_o;u,v) \leq \sum_{i=1}^{q} \frac{\partial\phi}{\partial y_i}(F(x_o)) \cdot f_i^{\circ\circ}(x_o;u,v)$$

$$+ \phi^{\circ\circ}(F(x_o);JF(x_o)\cdot u,JF(x_o)\cdot v).$$

De plus, si chaque f_i est deux fois continûment différentiable sur $\mathcal{O}$,

où, si Φ est deux fois continûment différentiable sur $\mathcal{O}'$ et si $q = 1$,

alors cette inégalité devient une égalité.

Preuve :

1°) Montrons que l'inégalité

$$(\phi \circ F)^{\circ\circ}(x_o;u,v) \leq \sum_{i=1}^{q} \frac{\partial\phi}{\partial y_i}(F(x_o)) \cdot f_i^{\circ\circ}(x_o;u,v)$$

$$+ \phi^{\circ\circ}(F(x_o);JF(x_o)\cdot u,JF(x_o)\cdot v)$$

est toujours vérifiée.

Soient u et v appartenant à $\mathbb{R}^p$.

Avant de commencer la démonstration proprement dite, nous allons établir une égalité qui sera utilisée par la suite.

D'une façon générale, notons par $\Theta(x,\lambda)$ une expression telle que

$$\lim_{\substack{x \to x_o \\ \lambda \to 0^+}} \Theta(x,\lambda) = 0.$$

Comme chaque composante f_i de F est de classe $\mathcal{C}^{1,1}$ sur $\mathcal{O}$, la fonction F est continûment différentiable sur $\mathcal{O}$. Ainsi,

$$F(x+\lambda u) = F(x) + \lambda\cdot JF(x_o)\cdot u + \lambda\cdot\Theta(x,\lambda). \tag{4.2}$$

D'où,

$$\nabla\phi(F(x+\lambda u)) = \nabla\phi(F(x) + \lambda\cdot JF(x_o)\cdot u) + \lambda\cdot\Theta(x,\lambda). \qquad (4.3)$$

En effet, pour x assez proche de x_o et λ assez petit, comme la fonction $\nabla\phi$ est Lipschitzienne de constante K dans un voisinage de $F(x_o)$, nous avons

$$0 \le \frac{\| \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x) + \lambda\cdot JF(x_o)\cdot u) \|}{\lambda}$$

$$\le \frac{K\cdot\| F(x+\lambda u) - F(x) - \lambda\cdot JF(x_o)\cdot u \|}{\lambda}.$$

Ainsi, en appliquant l'égalité (4.2), nous obtenons que

$$0 \le \frac{\| \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x) + \lambda\cdot JF(x_o)\cdot u) \|}{\lambda}$$

$$\le \frac{K\cdot\lambda\cdot\| \Theta(x,\varepsilon) \|}{\lambda}.$$

La définition de la fonction $\Theta(x,\lambda)$ entraîne alors que

$$\lim_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\| \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x) + \lambda\cdot JF(x_o)\cdot u) \|}{\lambda} = 0.$$

Autrement dit, l'égalité (4.3) est vraie.

Passons à la démonstration de l'inégalité du théorème.

Par la propriété 4.2.1, nous exprimons que

$$(\phi \circ F)^{\circ\circ}(x_o;u,v) = \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle\nabla(\Phi \circ F)(x+\lambda u) - \nabla(\phi \circ F)(x), v\rangle}{\lambda}.$$

Nous allons chercher à écrire l'expression concernée par la limite supérieure de telle façon qu'un passage à la limite supérieure fournisse le résultat voulu. Nous avons

$$\frac{\langle \nabla(\Phi \circ F)(x+\lambda u) - \nabla(\phi \circ F)(x) , v \rangle}{\lambda}$$

$$= \frac{\langle \nabla(\Phi \circ F)(x+\lambda u) , v \rangle}{\lambda} - \frac{\langle \nabla(\phi \circ F)(x) , v \rangle}{\lambda}.$$

Or, par les règles de différentiation, nous savons que

$$\nabla(\phi \circ F)(z) = [JF(z)]^T \cdot \nabla\phi(F(z))$$

pour tout z dans $\mathcal{O}$.

D'où,

$$\frac{\langle \nabla(\Phi \circ F)(x+\lambda u) - \nabla(\phi \circ F)(x) , v \rangle}{\lambda}$$

$$= \frac{\langle [JF(x+\lambda u)]^T \cdot \nabla\phi(F(x+\lambda u)) , v \rangle}{\lambda}$$

$$- \frac{\langle [JF(x)]^T \cdot \nabla\phi(F(x)) , v \rangle}{\lambda}.$$

Par un artifice de calcul, nous avons

$$\frac{\langle \nabla(\Phi \circ F)(x+\lambda u) - \nabla(\phi \circ F)(x) , v \rangle}{\lambda} \tag{4.4}$$

$$= \frac{\langle [JF(x+\lambda u) - JF(x)]^T \cdot \nabla\phi(F(x+\lambda u)) , v \rangle}{\lambda}$$

$$+ \frac{\langle [JF(x)]^T \cdot [\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x))] , v \rangle}{\lambda}.$$

Reformulons chacun des termes du membre de droite de cette égalité.

1.1°) Reformulons $\dfrac{\langle [JF(x+\lambda u) - JF(x)]^T \cdot \nabla\phi(F(x+\lambda u)) , v \rangle}{\lambda}$.

Commençons par écrire ce terme de manière différente.

$$\frac{\langle [JF(x+\lambda u) - JF(x)]^T \cdot \nabla\phi(F(x+\lambda u)) , v \rangle}{\lambda}$$

$$= \frac{\langle [JF(x+\lambda u) - JF(x)]^T \cdot \nabla\phi(F(x_o)) , v \rangle}{\lambda}$$

$$+ \frac{\langle [JF(x+\lambda u) - JF(x)]^T \cdot [\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x_o))] , v \rangle}{\lambda}.$$

Montrons que dans cette égalité, le dernier terme converge vers 0 si x tend vers x_o et si λ tend vers 0 par valeurs positives. Ainsi, nous pourrons l'écrire sous la forme $\Theta(x,\lambda)$.

L'inégalité de Cauchy-Schwarz indique que

$$0 \le \frac{| < [JF(x+\lambda u) - JF(x)]^T \cdot [\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x_o)] \ , \ v > |}{\lambda}$$

$$\le \frac{\| JF(x+\lambda u) - JF(x) \| \cdot \| \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x_o) \| \cdot \| v \|}{\lambda}.$$

Suite aux remarques *(R1)* et *(R4)* énoncées avant la propriété 4.2.4, prenons ε = mim $(\varepsilon_1, \varepsilon_F)$. Comme nous envisagerons que x tend vers x_o et que λ tend vers 0, nous pouvons supposer que x et $(x+\lambda u)$ appartiennent à $B(x_o, \varepsilon)$. D'où, nous avons que

$$0 \le \frac{| < [JF(x+\lambda u) - JF(x)]^T \cdot [\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x_o)] \ , \ v > |}{\lambda}$$

$$\le \frac{K_F \cdot \lambda \cdot \| u \| \cdot K_1 \cdot \| x - x_o + \lambda u \| \cdot \| v \|}{\lambda}.$$

L'inégalité triangulaire implique alors que

$$0 \le \frac{| < [JF(x+\lambda u) - JF(x)]^T \cdot [\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x_o)] \ , \ v > |}{\lambda}$$

$$\le K_F \cdot K_1 \cdot \| u \| \cdot \| v \| \cdot (\| x - x_o \| + \lambda \cdot \| u \|).$$

Et comme nous avons évidemment que

$$\lim_{\substack{x \to x_o \\ \lambda \to 0^+}} (\| x - x_o \| + \lambda \cdot \| u \|) = 0,$$

il suit que

$$\lim_{\substack{x \to x_o \\ \lambda \to 0^+}} \frac{|< [JF(x+\lambda u) - JF(x)]^T \cdot [\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x_o)] \ , \ v >|}{\lambda} = 0.$$

Donc, nous obtenons que l'égalité qui suit est vérifiée.

$$\frac{< [JF(x+\lambda u) - JF(x)]^T \cdot \nabla\phi(F(x+\lambda u)) \ , \ v >}{\lambda}$$

$$= \frac{< [JF(x+\lambda u) - JF(x)]^T \cdot \nabla\phi(F(x_o)) \ , \ v >}{\lambda} + \Theta(x,\lambda).$$

C'est-à-dire, en exprimant différement le produit intérieur du premier terme du second membre, que

$$\frac{< [JF(x+\lambda u) - JF(x)]^T \cdot \nabla\phi(F(x+\lambda u)) , v >}{\lambda} \tag{4.5}$$

$$= \sum_{i=1}^{q} \left[\frac{\partial\phi}{\partial y_i}(F(x_o)) \cdot \frac{< \nabla f_i(x+\lambda u) - \nabla f_i(x) , v >}{\lambda} \right]$$

$$+ \Theta(x,\lambda).$$

1.2°) Reformulons $\frac{< [JF(x)]^T \cdot [\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x))] , v >}{\lambda}$.

Nous avons évidemment l'égalité qui suit

$$\frac{< [JF(x)]^T \cdot [\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x))] , v >}{\lambda}$$

$$= \frac{< \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x)) , JF(x)\cdot v >}{\lambda}.$$

Ainsi, par un artifice de calcul,

$$\frac{< [JF(x)]^T \cdot [\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x))] , v >}{\lambda}$$

$$= \frac{< \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x)) , JF(x_o)\cdot v >}{\lambda}$$

$$+ \frac{< \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x)) , [JF(x) - JF(x_o)]\cdot v >}{\lambda}.$$

En utilisant l'égalité (4.3), nous obtenons

$$\frac{< [JF(x)]^T \cdot [\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x))] , v >}{\lambda}$$

$$= \frac{< \nabla\phi(F(x) + \lambda\cdot JF(x_o)\cdot u) - \nabla\phi(F(x)) , JF(x_o)\cdot v >}{\lambda}$$

$$+ < \Theta(x,\lambda) , JF(x_o)\cdot v >$$

$$+ \frac{<\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x)) , [JF(x) - JF(x_o)]\cdot v>}{\lambda}.$$

Montrons que les deux derniers termes de cette égalité convergent vers 0 si x tend vers x_o et si λ tend vers 0 par valeurs positives. Ainsi, nous pourrons écrire la somme de ces deux termes par $\Theta(x,\lambda)$.

1.2.1°) Montrons que $\lim\limits_{\substack{x \to x_o \\ \lambda \to o^+}} \langle\ \Theta(x,\lambda)\ ,\ JF(x_o)\cdot v\ \rangle = 0.$

Par définition de $\Theta(x,\lambda)$, cette limite est nécéssairement nulle.

1.2.2°) Montrons que

$$\lim_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x))\ ,\ [JF(x) - JF(x_o)]\cdot v\rangle}{\lambda} = 0.$$

L'inégalité de Cauchy-Schwarz indique que

$$0 \le \frac{\langle\ \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x))\ ,\ [JF(x) - JF(x_o)]\cdot v\ \rangle}{\lambda}$$

$$\le \frac{\|\ \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x))\ \|\cdot\|\ JF(x) - JF(x_o)\ \|\cdot\|\ v\ \|}{\lambda}.$$

Suite aux remarques *(R1)* et *(R4)* énoncées avant la propriété 4.2.4. Prenons ε = mim $(\varepsilon_1,\varepsilon_F)$. Comme nous envisagerons que x tend vers x_o et que λ tend vers 0, nous pouvons supposer que x et $(x+\lambda u)$ appartiennent à $B(x_o,\varepsilon)$. D'où, nous avons que

$$0 \le \frac{\langle\ \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x))\ ,\ [JF(x) - JF(x_o)]\cdot v\ \rangle}{\lambda}$$

$$\le \frac{K_1\cdot\lambda\cdot\|\ u\ \|\ \cdot\ K_F\cdot\|\ x - x_o\ \|\ \cdot\ \|\ v\ \|}{\lambda}.$$

C'est-à-dire que

$$0 \leq \frac{\langle \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x)) \ , \ [JF(x) - JF(x_o)]\cdot v \rangle}{\lambda}$$

$$\leq K_1 \cdot K_F \cdot \| u \| \cdot \| v \| \cdot \| x - x_o \| .$$

Nous savons que

$$\lim_{\substack{x \to x_o \\ \lambda \to o^+}} K_1 \cdot K_F \cdot \| u \| \cdot \| v \| \cdot \| x - x_o \| = 0.$$

Donc,

$$\lim_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle \nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x)) \ , \ [JF(x) - JF(x_o)]\cdot v \rangle}{\lambda} = 0.$$

Donc, nous pouvons écrire que

$$\frac{\langle \ [JF(x)]^T \cdot [\nabla\phi(F(x+\lambda u)) - \nabla\phi(F(x))] \ , \ v \ \rangle}{\lambda} \tag{4.6}$$

$$= \frac{\langle \nabla\phi(F(x) + \lambda \cdot JF(x_o)\cdot u) - \nabla\phi(F(x)) \ , \ JF(x_o)\cdot v \rangle}{\lambda} + \Theta(x,\lambda).$$

Finalement, compte tenu des reformulations (4.5) et (4.6), l'égalité (4.4) devient

$$\frac{\langle \nabla(\Phi \circ F)(x+\lambda u) - \nabla(\phi \circ F)(x) \ , \ v \rangle}{\lambda}$$

$$= \sum_{i=1}^{q} \left[\frac{\partial\phi}{\partial y_i}(F(x_o)) \cdot \frac{\langle \nabla f_i(x+\lambda u) - \nabla f_i(x) \ , \ v \rangle}{\lambda} \right]$$

$$+ \frac{\langle \nabla\phi(F(x) + \lambda \cdot JF(x_o)\cdot u) - \nabla\phi(F(x)) \ , \ JF(x_o)\cdot v \rangle}{\lambda}$$

$$+ \Theta(x,\lambda).$$

Sur cette dernière égalité, passons à la limite supérieure. D'où, nous obtenons

$$\limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle \nabla(\Phi \circ F)(x+\lambda u) - \nabla(\phi \circ F)(x) , v \rangle}{\lambda}$$

$$\leq \sum_{i=1}^{q} \left\{ \frac{\partial \phi}{\partial y_i}(F(x_o)) \cdot \left[\limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle \nabla f_i(x+\lambda u) - \nabla f_i(x) , v \rangle}{\lambda} \right] \right\}$$

$$+ \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle \nabla\phi(F(x) + \lambda \cdot JF(x_o) \cdot u) - \nabla\phi(F(x)) , JF(x_o) \cdot v \rangle}{\lambda},$$

car, par la propriété 1.3.5, la limite supérieure d'une somme est inférieure à la somme des limites supérieures.

Autrement dit, en tenant compte de la caractérisation 4.2.1 de la bifonction de support de la matrice Hessienne généralisée,

$$(\phi \circ F)^{\circ\circ}(x_o;u,v) \leq \sum_{i=1}^{q} \frac{\partial \phi}{\partial y_i}(F(x_o)) \cdot f_i^{\circ\circ}(x_o;u,v)$$

$$+ \phi^{\circ\circ}(F(x_o);JF(x_o) \cdot u, JF(x_o) \cdot v).$$

2°) Montrons que si chaque composante f_i de F est deux fois continûment différentiable sur $\mathcal{O}$,

où si la fonction ϕ est deux fois continûment différentiable sur $\mathcal{O}'$ et si $q = 1$,

alors,

$$(\phi \circ F)^{\circ\circ}(x_o;u,v) = \sum_{i=1}^{q} \frac{\partial \phi}{\partial y_i}(F(x_o)) \cdot f_i^{\circ\circ}(x_o;u,v)$$

$$+ \phi^{\circ\circ}(F(x_o);JF(x_o) \cdot u, JF(x_o) \cdot v).$$

Nous allons utiliser le résultat suivant.

(R) Soient ψ et φ deux fonctions définies sur $\mathbb{R}^p$ et à valeurs réelles et soit a un point de $\mathbb{R}^p$.

Supposons que $\lim\limits_{x \to a} \varphi(x)$ existe et vaut $\varphi(a)$

Alors $\lim\limits_{x \to a} \sup\, (\varphi + \psi)(x) = \lim\limits_{x \to a} \varphi(x) + \lim\limits_{x \to a} \sup\, \psi(x)$.

La preuve de cette propriété se trouver en annexe B.4.

2.1°) Supposons que chaque composante f_i de F est deux fois continûment différentiable sur $\mathcal{O}$.

Soit i compris entre 1 et q. Comme f_i est deux fois continûment différentiable en x_o, nous avons que la limite suivante existe, et nous pouvons calculer son expression par la propriété 1.5.4 de la dérivée directionnelle

$$\lim_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\nabla f_i(x+\lambda u) - \nabla f_i(x)}{\lambda} = \langle \nabla^2 f_i(x_o) \, , \, u \rangle.$$

Ainsi,

$$\lim_{\substack{x \to x_o \\ \lambda \to o^+}} \sup \frac{\langle \nabla f_i(x+\lambda u) - \nabla f_i(x) \, , \, v \rangle}{\lambda}$$

$$= \lim_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle \nabla f_i(x+\lambda u) - \nabla f_i(x) \, , \, v \rangle}{\lambda}.$$

D'où, compte tenu du résultat *(R)*, le point précédent devient

$$\limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle \nabla(\Phi \circ F)(x+\lambda u) - \nabla(\phi \circ F)(x) \ , \ v \rangle}{\lambda}$$

$$= \sum_{i=1}^{q} \left\{ \frac{\partial\phi}{\partial y_i}(F(x_o)) \cdot \left[\lim_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle \nabla f_i(x+\lambda u) - \nabla f_i(x) \ , \ v \rangle}{\lambda} \right] \right\}$$

$$+ \limsup_{\substack{x \to x_o \\ \lambda \to o^+}} \frac{\langle \nabla\phi(F(x) + \lambda \cdot JF(x_o) \cdot u) - \nabla\phi(F(x)) \ , \ JF(x_o) \cdot v \rangle}{\lambda}.$$

C'est-à-dire que

$$(\phi \circ F)^{\circ\circ}(x_o;u,v) = \sum_{i=1}^{q} \frac{\partial\phi}{\partial y_i}(F(x_o)) \cdot \ f_i^{\circ\circ}(x_o;u,v)$$

$$+ \phi^{\circ\circ}(F(x_o);JF(x_o) \cdot u, JF(x_o) \cdot v).$$

2.2°) Supposons que la fonction ϕ est deux fois continûment différentiable sur $\mathcal{O}'$ et que $q = 1$.

De manière similaire au point point 2.1, nous montrons que l'égalité est vraie si les hypothèses citées ci-dessus sont vérifiées.

□

Par la propriété 4.1.5, nous savons que les ensembles $\partial^2(\phi \circ F)(x_o) \cdot u$, $\partial^2 f_i(x_o) \cdot u$ et $\partial^2\phi(F(x_o)) \cdot u$ sont convexes et fermés. Ainsi, compte tenu de la définition de support de chacun de ces ensembles, la propriété 1.7.4 entraîne que le théorème 4.2.5 peut être reformulé de la façon suivante.

Théorème 4.2.6 : Pour tout u appartenant à $\mathbb{R}^p$,

$$\partial^2(\phi \circ F)(x_o)\cdot u \subseteq \sum_{i=1}^{q} \left[\frac{\partial\phi}{\partial y_i}(F(x_o)) \cdot \partial^2 f_i(x_o)\cdot u \right]$$

$$+ [JF(x_o)]^T \cdot \partial^2\phi(F(x_o))\cdot JF(x_o)\cdot u.$$

De plus, si chaque composante f_i de F est deux fois continûment différentiable sur $\mathcal{O}$,

où si la fonction ϕ est deux fois continûment différentiable sur $\mathcal{O}'$ et si q = 1,

alors cette inclusion devient une égalité.

A présent illustrons ces deux théorèmes en développant quelques exemples.

Exemple 4.2.7 :

Considérons la fonction F définie sur $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}^q$ par

$$F(x) = A\cdot x + b$$

pour tout x dans $\mathbb{R}^p$, où A est une matrice de $\mathcal{M}(q,p)$ et b est un point de $\mathbb{R}^q$. Comme cette fonction est linéaire, elle est deux fois continûment différentiable sur $\mathbb{R}^p$.

Soient ϕ une fonction de classe $C^{1,1}$ définie sur $\mathbb{R}^q$ et à valeurs réelles et x_o appartenant à $\mathbb{R}^p$.

Ainsi, le théorème 4.2.5 indique pour tout points u et v de $\mathbb{R}^p$, que

$$(\phi \circ F)^{\circ\circ}(x_o;u,v) = \sum_{i=1}^{q} \frac{\partial\phi}{\partial y_i}(F(x_o)) \cdot f_i^{\circ\circ}(x_o;u,v)$$
$$+ \phi^{\circ\circ}(F(x_o);JF(x_o)\cdot u, JF(x_o)\cdot v).$$

Considérons alors u et v deux points de $\mathbb{R}^p$.

Comme la fonction F est deux fois continûment différentiable sur $\mathbb{R}^p$, chacune de ses composantes f_i l'est également.

Ainsi, grâce à la propriété 4.1.8, nous avons que

$$\partial^2 f_i(x_o) = \{ \nabla^2 f_i(x_o) \},$$

et nous calculons que

$$\nabla^2 f_i(x_o) = \nabla^2(A\cdot x_o + b)_i = \nabla(A)_i = 0 \text{ (dans } \mathcal{M}(p,p)),$$

où $(A\cdot x_o + b)_i$ désigne la $i^{ème}$ composante de $(A\cdot x_o + b)$, et où $(A)_i$ désigne la $i^{ème}$ ligne de la matrice A.

D'où, $\partial^2 f_i(x_o) = \{ 0 \}$ (dans $\mathcal{M}(p,p)$), donc $f_i^{\circ\circ}(x_o;u,v) = 0$.

Par linéarité de la fonction F, nous calculons également que

$$JF(x_o) = A.$$

Suite à ces calculs, l'égalité du théorème 4.2.5 devient

$$(\phi \circ F)^{\circ\circ}(x_o;u,v) = \phi^{\circ\circ}(A\cdot x_o + b; A\cdot u, A\cdot v).$$

Et par le théoréme 4.2.6, nous obtenons que

$$\partial^2(\phi \circ F)(x_o)\cdot u = A^T\cdot\partial^2\phi(F(x_o))\cdot A\cdot u.$$

■

Exemple 4.2.8 :

Soit F une fonction définie sur $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}^q$ telle que chacune de ses composantes soit de classe $C^{1,1}$ sur $\mathbb{R}^p$.

Soit la fonction ϕ définie sur $\mathbb{R}^q$ et à valeurs réelles par

$$\Phi(y) = \langle a,y\rangle$$

pour tout y dans $\mathbb{R}^q$, où a est un point de $\mathbb{R}^q$. Comme cette fonction est linéaire, elle est deux fois continûment différentiable sur $\mathbb{R}^q$. Soit x_o un point de $\mathbb{R}^p$.

Ainsi, le théorème 4.2.5 indique pour tous les points u et v de $\mathbb{R}^p$, que

$$(\phi \circ F)^{\circ\circ}(x_o;u,v) \le \sum_{i=1}^{q} \frac{\partial\phi}{\partial y_i}(F(x_o)) \cdot f_i^{\circ\circ}(x_o;u,v)$$

$$+ \phi^{\circ\circ}(F(x_o);JF(x_o)\cdot u,JF(x_o)\cdot v).$$

Considérons alors u et v deux points de $\mathbb{R}^p$.

Comme la fonction ϕ est deux fois différentiable en $F(x_o)$, la propriété 4.1.8 entraîne que

$$\partial^2\phi(F(x_o)) = \{ \nabla^2\Phi(f(x_o)) \}.$$

Et comme la fonction ϕ est linéaire, nous calculons que

$$\nabla^2\Phi(f(x_o)) = 0 \text{ (dans } \mathcal{M}(q,q)).$$

Finalement, il suit de ces deux égalités que

$$\partial^2\phi(F(x_o)) = \{ 0 \} \text{ (dans } \mathcal{M}(q,q)).$$

De plus, par définition de la fonction ϕ, nous calculons que

$$\frac{\partial\phi}{\partial y_i}(F(x_o)) = \frac{\partial \langle a,F(x_o)\rangle}{\partial y_i} = \frac{\partial \sum_{j=1}^{q} a_j \cdot f_j(x_o)}{\partial y_i} = a_i,$$

où a_j désigne la $j^{\text{ème}}$ composante du vecteur a.

Comme nous avons l'égalité $(\phi \circ F)(x) = \langle a,F(x)\rangle$ pour tout x dans $\mathbb{R}^p$, nous noterons la fonction $(\phi \circ F)$ par $\langle a,F\rangle$.

Suite à ces constatations, l'inégalité du théorème 4.2.5 devient

$$\langle a,F\rangle^{\circ\circ}(x_o;u,v) \leq \sum_{i=1}^{q} a_i \cdot f_i^{\circ\circ}(x_o;u,v).$$

D'où, par application du théorème 4.2.6,

$$\partial^2\langle a,F\rangle(x_o)\cdot u \subseteq \sum_{i=1}^{q} a_i \cdot \partial^2 f_i(x_o)\cdot u.$$

■

Exemple 4.2.9 :

Soit f une fonction de classe $C^{1,1}$ définie sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles.

Soit ϕ une fonction deux fois continûment différentiable définie sur l'intervalle réel ouvert I et à valeurs réelles.

Supposons que $f(\mathcal{O})$ soit inclus dans I.

Comme la fonction f est à valeurs réelles, nous noterons JF(x) par $[\nabla f(x)]^T$ quel que soit x dans $\mathcal{O}$. Et comme ϕ est définie sur un intervalle réel, nous noterons $\nabla\phi(y)$ par $\Phi'(y)$ et $\nabla^2\Phi(y)$ par $\Phi''(y)$ quel que soit y dans I.

Ainsi, le théorème 4.2.5 indique que, pour tout u et v dans $\mathbb{R}^p$,

$$(\phi \circ f)^{\circ\circ}(x_o;u,v) = \phi'(f(x_o))\cdot f^{\circ\circ}(x_o;u,v)$$
$$+ \phi^{\circ\circ}(f(x_o);[\nabla f(x_o)]^T\cdot u,[\nabla f(x_o)]^T\cdot v).$$

Considérons alors u et v deux points de $\mathbb{R}^p$.

Mais, par définition de la bifonction de support, nous savons que

$$\phi^{\circ\circ}(f(x_o);[\nabla f(x_o)]^T\cdot u,[\nabla f(x_o)]^T\cdot v)$$
$$= \max \{ \langle M\cdot[\nabla f(x_o)]^T\cdot u,[\nabla f(x_o)]^T\cdot v\rangle \mid M \in \partial^2\phi(f(x_o)) \}.$$

Or, la fonction ϕ étant deux fois continûment différentiable sur l'intervalle I, il suit de la propriété 4.1.8 que

$$\partial^2\phi(f(x_o)) = \{ \phi''(f(x_o)) \}.$$

D'où,

$$\phi^{\circ\circ}(f(x_o);[\nabla f(x_o)]^T\cdot u,[\nabla f(x_o)]^T\cdot v)$$
$$= \langle\phi''(f(x_o))\cdot[\nabla f(x_o)]^T\cdot u,[\nabla f(x_o)]^T\cdot v\rangle.$$

C'est-à-dire que

$$\phi^{\circ\circ}(f(x_o);[\nabla f(x_o)]^T\cdot u,[\nabla f(x_o)]^T\cdot v)$$
$$= \phi''(f(x_o))\cdot\langle\nabla f(x_o),u\rangle\cdot\langle\nabla f(x_o),v\rangle.$$

Finalement, grâce à cette égalité, nous obtenons que

$$(\phi \circ f)^{\circ\circ}(x_o;u,v) = \phi'(f(x_o))\cdot f^{\circ\circ}(x_o;u,v)$$
$$+ \phi''(f(x_o))\cdot\langle\nabla f(x_o),u\rangle\cdot\langle\nabla f(x_o),v\rangle.$$

Ainsi, suite au théorème 4.2.6, nous avons

$$\partial^2(\phi \circ F)(x_o)\cdot u = \phi'(f(x_o))\cdot\partial^2 f(x_o)\cdot u$$
$$+ \phi''(f(x_o))\cdot\| \nabla f(x_o) \|^2\cdot u.$$

■

4.3 Extension du dévelopement de Taylor d'ordre un pour des fonctions de classe $C^{1,1}$.

Pour une fonction f de classe $C^{1,1}$ sur $\mathcal{O}$, nous allons utiliser la matrice Hessienne généralisée de f en point x de $\mathcal{O}$ afin de déterminer une généralisation du développement de Taylor du premier ordre de la fonction f.

Commençons par généraliser le théorème de Rolle.

Théorème 4.3.1 : Soit ξ une fonction définie sur $\mathbb{R}$ et à valeurs réelles.

Supposons que la fonction ξ soit localement Lipschitzienne sur l'intervalle réel $[a,b]$ et que $\xi(a) = \xi(b)$,

alors il existe un point c de l'intervalle ouvert $]a,b[$ tel que 0 appartient à $\partial\xi(c)$.

Preuve :

Remarquons que la fonction ξ est continue sur $[a,b]$ car elle est localement Lipschitzienne sur l'intervalle réel $[a,b]$. Ainsi, elle atteint ses bornes en des points de l'intervalle compact $[a,b]$.
D'où, nous pouvons considérer les points u et v de $[a,b]$ tels que

$$\xi(u) = \inf \{ \xi(x) \mid x \in [a,b] \}, \text{ et}$$

$$\xi(v) = \sup \{ \xi(x) \mid x \in [a,b] \}.$$

- Si $\xi(u) = \xi(v)$, alors la fonction ξ est constante sur $[a,b]$. D'où, elle est dérivable sur l'intervalle $]a,b[$, et $\xi'(c) = 0$ pour tout c dans $]a,b[$. De plus, nous savons que $\partial\xi(c) = \{\xi'(c)\}$ pour tout c dans $]a,b[$. Ainsi, $\partial\xi(c) = \{ 0 \}$ pour tout c dans $]a,b[$. Donc pour n'importe quel point c de $]a,b[$, nous avons que 0 appartient à $\partial\xi(c)$.
- Si $\xi(u) < \xi(v)$, alors au moins un des points u et v est intérieur à $[a,b]$.

  En effet, si ce n'est pas le cas, alors $\xi(u) = \xi(v)$ car $\xi(a) = \xi(b)$.

Notons par c ce point intérieur. Par définition des points u et v, le point c est un point extrémum de la fonction ξ sur $]a,b[$. D'où, par application de la condition 2.2.7 d'extrémum, le point c est tel que 0 appartient à $\partial\xi(c)$.

□

Nous allons utiliser ce théorème pour démontrer la propriété qui suit.

Lemme 4.3.2 : Soit I un intervalle réel contenant $[0,1]$.
Considérons une fonction ϕ de classe $C^{1,1}$ sur I,
alors il existe un point c de $[0,1]$ tel que

$$\phi(1) - \phi(0) - \phi'(0) \text{ appartient à } \frac{1}{2}\cdot\partial^2\phi(c).$$

Preuve :

Considérons la fonction auxiliaire α définie sur l'intervalle $[0,1]$ et à valeurs réelles par

$$\alpha(t) = \phi(t) + \phi'(t)\cdot(1-t) + \frac{R}{2}\cdot(1-t)^2,$$

pour tout t dans $[0,1]$, où R est une constante.

Nous allons envisager $\partial\alpha(t)$. Mais auparavant, il faut montrer que la fonction α est localement Lipschitzienne sur $[0,1]$.

1°) Montrons que la fonction α est localement Lipschitzienne sur $[0,1]$.

Soit t_0 un point de $[0,1]$. Il faut trouver des constantes $\varepsilon > 0$ et $K > 0$ telles que $|\alpha(s) - \alpha(t)| \leq K\cdot|s - t|$ pour tout s et t dans l'intervalle $]t_0-\varepsilon, t_0+\varepsilon[$.

Remarquons d'abord que, comme ϕ appartient à $C^{1,1}(I)$, les points suivants sont vrais.

- ϕ est une fonction localement Lipschitzienne en t_o. Autrement dit, il existe des constantes $\eta > 0$ et $L > 0$ telles que

  $$| \phi(s) - \phi(t) | \leq L \cdot | s - t |$$

  pour tout s et t dans l'intervalle $]t_o - \eta, t_o + \eta[$.

- ϕ' est une fonction localement Lipschitzienne en t_o. Autrement dit, il existe des constantes $\theta > 0$ et $M > 0$ telles que

  $$| \phi'(s) - \phi'(t) | \leq M \cdot | s - t |$$

  pour tout s et t dans l'intervalle $]t_o - \theta, t_o + \theta[$.

- ϕ' est une fonction continue en tout point t de $[0,1]$. Ainsi, il existe une constante $C > 0$ telle que $| \phi'(t) | \leq C$ pour tout t dans $[0,1]$.

Cherchons les constantes ε et K.

Par définition de la fonction α,

$$| \alpha(s) - \alpha(t) | = | \phi(s) + \phi'(s) \cdot (1-s) + \frac{R}{2} \cdot (1-s)^2 - \phi(t) - \phi'(t) \cdot (1-t) - \frac{R}{2} \cdot (1-t)^2 |$$

L'inégalité triangulaire entraîne alors que

$$| \alpha(s) - \alpha(t) | = | \phi(s) - \phi(t) | + | \phi'(s) - \phi'(t) | + | \phi'(t) | \cdot | s - t | + \frac{R}{2} \cdot | (1-s)^2 - (1-t)^2 | .$$

Or, nous calculons que

$$| (1-s)^2 - (1-t)^2 | = | 2 \cdot (t - s) + (t + s) \cdot (s - t) | \leq 2 \cdot | s - t | + | t + s | \cdot | s - t | \leq 4 \cdot | s - t | ,$$

car s et t appartiennent à $[0,1]$.

D'où,

$$| \alpha(s) - \alpha(t) | = | \phi(s) - \phi(t) | + | \phi'(s) - \phi'(t) |$$
$$+ | \phi'(t) | \cdot | s - t | + 2 \cdot R \cdot | s - t | .$$

Choisissons $\varepsilon = \min(\eta, \theta)$. Ainsi, avec s et t appartenant à $]t_o - \varepsilon, t_o + \varepsilon[$,

$$| \alpha(s) - \alpha(t) | \leq (L + M + C + 2 \cdot R) \cdot | s - t | .$$

Donc, en prenant $K = (L + M + C + 2 \cdot R)$, nous obtenons que α est une fonction localement Lipschitzienne en t_o.

2°) Montrons qu'il existe un point c de $[0,1]$ tel que

$\phi(1) - \phi(0) - \phi'(0)$ appartient à $\frac{1}{2} \cdot \partial^2 \phi(c)$.

Soit t un point de $[0,1]$.

Par définition de la fonction α, nous avons

$$\partial\alpha(t) = \partial[\phi(t) + \phi'(t) \cdot (1-t) + \frac{R}{2} \cdot (1-t)^2].$$

D'où, par les propriétés 2.2.5 et 2.2.6, nous obtenons

$$\partial\alpha(t) \subseteq \partial\phi(t) + \partial[\phi'(t) \cdot (1-t)] + \frac{R}{2} \cdot \partial[(1-t)^2].$$

Or, nous calculons que

$\partial\phi(t) = \{ \phi'(t) \}$ car ϕ est continûment différentiable sur l'intervalle $[0,1]$.

$\partial[\phi'(t) \cdot (1-t)] = \partial^2\phi(t) \cdot (1-t) - \{ \phi'(t) \}$ car ϕ appartient à $\mathcal{C}^{1,1}(I)$ et la fonction $(1-t)$ est continûment différentiable en tout point t de $[0,1]$.

$\partial[(1-t)^2] = \{ -2 \cdot (1-t) \}$ car la fonction $(1-t)^2$ est continûment différentiable en tout point t de $[0,1]$.

Nous obtenons donc que

$$\partial\alpha(t) \subseteq \partial^2\phi(t)\cdot(1-t) - R\cdot(1-t).$$

C'est-à-dire que

$$\partial\alpha(t) \subseteq [\partial^2\phi(t) - R]\cdot(1-t).$$

Prenons R tel que $\alpha(0) = \alpha(1)$. Ainsi, le théorème 4.3.1 entraîne qu'il existe un point c de l'intervalle $]0,1[$ tel que

$$0 \in \partial\alpha(c).$$

D'où,

$$0 \in [\partial^2\phi(c) - R]\cdot(1-c).$$

Comme $(1-c)$ est un nombre constant non nul, il suit que

$$0 \in [\partial^2\phi(c) - R].$$

Autrement dit,

$$R \in \partial^2\phi(c).$$

De plus, comme nous avons que $\alpha(0) = \alpha(1)$, la définition de la fonction α entraîne que

$$\phi(1) = \phi(0) + \phi'(0) + \frac{R}{2}.$$

Donc, nous obtenons finalement que

$$\phi(1) - \phi(0) - \phi'(0) \in \frac{1}{2}\cdot\partial^2\phi(c).$$

En conclusion, nous venons de trouver un point c de $]0,1[$ tel que

$\phi(1) - \phi(0) - \phi'(0)$ appartient à $\frac{1}{2}\cdot\partial^2\phi(c)$.

□

De cette propriété, nous déduisons le théorème suivant.

Théorème 4.3.3 : Soit f une fonction de classe $\mathcal{C}^{1,1}$ sur $\mathcal{O}$ et soit $[a,b]$ inclus dans $\mathcal{O}$,

alors il existe un point c de $]a,b[$ et une matrice M_c de $\partial^2 f(c)$ tels que

$$f(b) = f(a) + \langle \nabla f(a), b - a\rangle + \frac{1}{2}\cdot(b-a)^T \cdot M_c \cdot (b-a).$$

Preuve :

Considérons la fonction auxiliaire k définie sur l'intervalle $[0,1]$ et à valeurs dans $\mathbb{R}^p$ par $k(t) = (a + t\cdot(b-a))$ pour tout t dans $[0,1]$. Comme elle est linéaire, elle est deux fois continûment différentiable sur $[0,1]$. D'où, elle est de classe $\mathcal{C}^{1,1}$ sur $[0,1]$.

Construisons alors la fonction ϕ définie sur l'intervalle $[0,1]$ et à valeurs réelles par $\phi(t) = (f \circ k)(t)$ pour tout t dans $[0,1]$. Par application de la propriété 4.2.4, cette fonction ϕ est de classe $\mathcal{C}^{1,1}$ sur $[0,1]$. De plus, la propriété 4.3.2 indique qu'il existe un point d appartenant à $]0,1[$ tel que

$$\phi(1) - \phi(0) - \phi'(0) \in \frac{1}{2}\cdot \partial^2\phi(d).$$

Autrement dit, il existe un point d de $]0,1[$ et un point m_d de $\partial^2\phi(d)$ tels que

$$\phi(1) - \phi(0) - \phi'(0) = \frac{1}{2}\cdot m_d. \tag{4.7}$$

Or, par définition de la fonction ϕ, nous calculons que

- $\phi(1) = f(k(1)) = f(b)$,
- $\phi(0) = f(k(0)) = f(a)$, et
- $\phi'(0) = \langle \nabla f(k(0)), k'(0) \rangle = \langle \nabla f(a), b - a \rangle$.

Et, suite à l'exemple 4.2.7, nous savons que, pout tout t appartenant à [0,1],

$$\partial^2\phi(t) = (b - a)^T \cdot \partial^2 f(a + t\cdot(b - a))\cdot(b - a).$$

Posons c = a + d·(b − a) (comme d appartient à]0,1[, il suit que c est dans]a,b[). D'où,

$$\partial^2\phi(d) = (b - a)^T \cdot \partial^2 f(c)\cdot(b - a).$$

Grâce à cette égalité, nous savons qu'il existe un matrice M_c de $\partial^2 f(c)$ telle que

$$m_d = (b - a)^T \cdot M_c \cdot (b - a).$$

Ainsi, ces remarques impliquent que l'égalité (4.7) devient

$$f(b) - f(a) - \langle \nabla f(a), b - a \rangle = \frac{1}{2}\cdot(b - a)^T \cdot M_c \cdot (b - a).$$

Autrement dit, nous obtenons finalement que

$$f(b) = f(a) + \langle \nabla f(a), b - a \rangle + \frac{1}{2}\cdot(b - a)^T \cdot M_c \cdot (b - a).$$

□

Pour des fonctions de classes $\mathcal{C}^{1,1}$, ce théorème nous fournit une inégalité qui étend le développement de Taylor d'ordre un existant dans le cas deux fois continûment différentiable.

Cependant, le théorème assure uniquement l'existence d'une matrice vérifiant l'égalité, mais il ne permet pas de déterminer cette matrice.

Néanmoins, cette extension rend possible les deux propriétés suivantes qui caractérisent la convexité d'une foncion de classe $\mathcal{C}^{1,1}$.

Propriété 4.3.4 : Soit f une fonction appartenant à $C^{1,1}(\mathcal{O})$, alors, f est convexe sur $\mathcal{O}$ si et seulement si l'ensemble $\partial^2 f(x)$ est semi-défini positif pour tout x dans $\mathcal{O}$.

Preuve :

1°) Montrons que si f est convexe sur $\mathcal{O}$,

alors $\partial^2 f(x)$ en semi-défini positif.

Soit x dans $\mathcal{O}$ et M une matrice de $\partial^2 f(x)$.

Il faut montrer que $y^T \cdot M \cdot y \geq 0$ pour tout y dans $\mathbb{R}^p$.

Par définition de $\partial^2 f(x)$, il existe un nombre k de $\mathbb{N}_0$ tel que $M = \sum_{i=1}^{k} \lambda_i \cdot M_i$, où, pour tout i compris entre 1 et k, M_i appartient à $\partial^2(f,x)$, $\lambda_i \geq 0$ et $\sum_{i=1}^{k} \lambda_i = 1$.

Comme M_i est une matrice de $\partial^2(f,x)$, il existe une suite $(x_n^i)_{n=1}^{\infty}$ de points de $\Omega^c_{\nabla f}$ qui converge vers x telle que $M_i = \lim_{n\to\infty} \nabla^2 f(x_n^i)$.

Soit y un point de $\mathbb{R}^p$.

Ainsi, $y^T \cdot M_i \cdot y = \lim_{n\to\infty} y^T \cdot \nabla^2 f(x_n^i) \cdot y$.

Or, nous savons que $y^T \cdot \nabla^2 f(x_n^i) \cdot y \geq 0$ pour tout $n \geq 1$ car $\nabla^2 f(x_n^i)$ est une matrice semi-définie positive.

D'ou, par passage à la limite $y^T \cdot M_i \cdot y \geq 0$. Et donc, $y^T \cdot M \cdot y \geq 0$.

Autrement dit, $\partial^2 f(x)$ est semi-défini positif.

2°) Montrons que si $\partial^2 f(x)$ est semi-défini positif pour tout x dans $\mathcal{O}$,

alors f est convexe sur $\mathcal{O}$.

La propriété 1.7.10 indique qu'il suffit de voir que

$$< \nabla f(x) - \nabla f(y) \, , \, x - y > \geq 0$$

pour tout x et y dans $\mathcal{O}$.

Soient x et y deux points de $\mathcal{O}$.

Considérons le résultat du théorème 4.3.3. Il entraîne qu'il existe un point c de]x,y[et une matrice M_c de $\partial^2 f(c)$ tels que

$$f(y) = f(x) + < \nabla f(x) \, , \, y - x > + \frac{1}{2} \cdot (y - x)^T \cdot M_c \cdot (y - x).$$

Comme la matrice M_c est semi-définie positive, nous savons que

$$(y - x)^T \cdot M_c \cdot (y - x) \geq 0.$$

D'où,

$$f(y) \geq f(x) + < \nabla f(x) \, , \, y - x >. \tag{4.8}$$

De manière identique, nous obtenons, en inversant les rôles de x et y, que

$$f(x) \geq f(y) + < \nabla f(y) \, , \, x - y >,$$

c'est-à-dire que

$$-f(y) \geq -f(x) - < \nabla f(y) \, , \, y - x >. \tag{4.9}$$

Sommons les inégalités (4.8) et (4.9).

Ainsi,

$$0 \geq < \nabla f(x) - \nabla f(y) \, , \, y - x >.$$

Autrment dit,

$$0 \leq < \nabla f(x) - \nabla f(y) \, , \, x - y >.$$

□

Propriété 4.3.5 : Soit f une fonction de classe $C^{1,1}$ sur $\mathcal{O}$,
alors pour que f soit strictement convexe sur $\mathcal{O}$,
il suffit que $\partial^2 f(x)$ soit défini positif
pour tout x dans $\mathcal{O}$.

La démonstration se déroule selon le même schéma que le point deux de la propriété précédente. C'est pourquoi nous ne la développerons pas.

4.4 CONDITIONS D'OPTIMALITÉ DU SECOND ORDRE POUR DES PROBLÈMES DÉCRITS PAR DES FONCTIONS DE CLASSE $C^{1,1}$.

Nous allons établir des conditions nécéssaires du second ordre pour des problèmes où les fonctions sont de classe $C^{1,1}$. Nous séparerons les problèmes sans contrainte des problèmes avec contraintes.

4.4.A: Le cas des problèmes sans contraintes.

Soit f une fonction de classe $C^{1,1}$ sur $\mathcal{O}$. La forme générale du problème est

$$(PSC) \left\{ \text{Minimiser } f(x) \text{ pour tout } x \text{ dans } \mathcal{O}. \right.$$

Le théorème suivant fournit des conditions recherchées.

Théorème 4.4.1 : Soit x_o appartenant à $\mathcal{O}$,

si x_o est un point de minimum local pour le problème *(PSC)*,

alors pour tout d de $\mathbb{R}^p$, il existe une matrice A de $\partial^2 f(x_o)$ telle que $d^T \cdot A \cdot d \geq 0$.

Preuve :

Soit d appartenant à $\mathbb{R}^p$.

Considérons la suite $(x_n)_{n=1}^{\infty}$ définie par $x_n = (x_o + \frac{1}{n} \cdot d)$ pour tout $n \geq 1$. Sans perte de généralité, nous pouvons supposer que la suite $(x_n)_{n=1}^{\infty}$ est incluse dans $\mathcal{O}$, sinon il suffit de négliger les premiers termes de la suite.

Soit $n \geq 1$. Appliquons le théorème 4.3.3 à la fonction f sur l'intervalle $[x_o, x_n]$. Il nous indique qu'il existe ξ_n dans $]x_o, x_n[$ et A_n appartenant à $\partial^2 f(\xi_n)$ tels que

$$f(x_n) = f(x_o) + \langle \nabla f(x_o) \, , \, x_n - x_o \rangle + \frac{1}{2} \cdot (x_n - x_o)^T \cdot A_n \cdot (x_n - x_o).$$

Or, par définition de la suite $(x_n)_{n=1}^{\infty}$, nous calculons que

$$x_n - x_o = \frac{1}{n} \cdot d.$$

D'où,

$$f(x_n) = f(x_o) + \frac{1}{n} \cdot \langle \nabla f(x_o) \, , \, d \rangle + \frac{1}{2 \cdot n^2} \cdot d^T \cdot A_n \cdot d.$$

Le point x_o étant un minimum local pour le problème *(PSC)*, il suit que $\nabla f(x_o) = 0$ et $f(x_n) \geq f(x_o)$. D'où,

$$d^T \cdot A_n \cdot d \geq 0.$$

Ainsi, nous obtenons une suite $(A_n)_{n=1}^{\infty}$ de matrices telles que A_n appartient à $\partial^2 f(\xi_n)$ et $d^T \cdot A_n \cdot d \geq 0$.

Notons que la suite $(\xi_n)_{n=1}^{\infty}$ converge vers x_o car ξ_n appartient à $]x_o,x_n[$ et la suite $(x_n)_{n=1}^{\infty}$ converge vers x_o.
De plus, grâce à la propriété 4.1.6, nous savons que l'application $\partial^2 f$ est localement bornée. D'où, la suite $(A_n)_{n=1}^{\infty}$ est bornée. Le théorème de Bolzano-Weierstrass entraîne alors qu'il existe une sous-suite $(A_{q_n})_{n=1}^{\infty}$ extraite de $(A_n)_{n=1}^{\infty}$ qui converge vers une matrice A de $\mathcal{M}(p,p)$.
Finalement, nous avons une suite $(\xi_{q_n})_{n=1}^{\infty}$ convergente vers x_o et une suite $(A_{q_n})_{n=1}^{\infty}$ convergente vers une matrice A de $\mathcal{M}(p,p)$. La propriété 4.1.7 indiquant que l'application $\partial^2 f$ est semi-continue supérieurement, nous obtenons que A appartient à $\partial^2 f(x_o)$.
De plus, comme $d^T \cdot A_n \cdot d \geq 0$ pour tout $n \geq 1$, il suit, par passage à la limite, que $d^T \cdot A \cdot d \geq 0$.

□

Ce théorème appele à quelques commentaires que nous allons développer maintenant.

Commentaire 4.4.2

Pour tout point d de $\mathbb{R}^p$, le théorème 4.4.1 assure qu'il existe une matice A de $\partial^2 f(x_o)$ telle que $d^T \cdot A \cdot d \geq 0$. Cependant, cette inégalité ne peut être étendue à toutes les matrices de $\partial^2 f(x_o)$.
Le contre-exemple que nous allons décrire ici illustre cette remarque.

Soit f une fonction définie sur $\mathbb{R}$ et à valeurs réelles par

$$f(x) = \begin{cases} \int_0^x \phi(t)\, dt & \text{si } x \geq 0 \\ \int_0^{-x} \phi(t)\, dt & \text{si } x < 0 \end{cases},$$

$$\text{où } \phi(t) = \begin{cases} 2\cdot t^2 + t^2 \cdot \sin(\frac{1}{t}) & \text{si } t > 0 \\ 0 & \text{si } t = 0 \end{cases}.$$

1°) Nous allons voir que f est de classe $C^{1,1}$ sur $\mathbb{R}$.

Nous allons utiliser les fonctions suivantes.

$f_+(x) = \int_0^x \phi(t)\, dt$ définie sur $\mathbb{R}_0^+$, et

$f_-(x) = \int_0^{-x} \phi(t)\, dt$ définie sur $\mathbb{R}_0^-$.

Comme $\lim_{t\to 0} 2\cdot t^2 + t^2\cdot \sin(\frac{1}{t}) = 0$, la fonction ϕ est continue pour tout $t \geq 0$. De plus, si ϕ est également considérée comme fonction de x, elle est dérivable par rapport à x. Et nous calculons que $\phi_x'(t) = 0$ pour tout $t \geq 0$. Ainsi, la fonction ϕ_x' est une fonction continue. D'où, grâce aux propriétés 1.10.2 et 1.10.3, il suit que les fonction f_+ et f_- sont continues et dérivables sur leur domaine respectif.

De plus, la propriété 1.10.3 fournit l'expression de chacune desdérivées de ces fonctions. C'est-à-dire que

$f_+'(x) = 2\cdot x^2 + x^2\cdot \sin(\frac{1}{x})$, et

$f_-'(x) = 2\cdot x^2 - x^2\cdot \sin(\frac{1}{x})$.

Ce qui précède veut dire que f est dérivable pour tout x non nul.

Nous allons voir que f est dérivable sur $\mathbb{R}$ et que

$$f'(x) = \begin{cases} 2\cdot x^2 + x^2\cdot \sin(\frac{1}{x}) & \text{si } x > 0 \\ 0 & \text{si } x = 0. \\ 2\cdot x^2 - x^2\cdot \sin(\frac{1}{x}) & \text{si } x < 0 \end{cases}$$

Il suffit démontrer que f est dérivable en 0 et que $f'(0) = 0$.

Autrement dit, il faut voir que

$$\lim_{x\to 0^+} \frac{f(x) - f(0)}{x - 0} = 0 \text{ et que } \lim_{x\to 0^-} \frac{f(x) - f(0)}{x - 0} = 0.$$

Montrons que $\lim_{x\to 0^+} \frac{f(x) - f(0)}{x - 0} = 0$.

Comme $x > 0$, nous savons que $f(x) = \int_0^x 2\cdot t^2 + t^2\cdot \sin(\frac{1}{t})\ dt$.

Nous savons également que $f(0) = 0$.

Ainsi, $\frac{f(x) - f(0)}{x - 0} = \frac{\int_0^x 2\cdot t^2 + t^2\cdot \sin(\frac{1}{t})\ dt}{x}$.

Or, $t^2 \le 2\cdot t^2 + t^2\cdot \sin(\frac{1}{t}) \le 3\cdot t^2$, d'où

$$\int_0^x t^2\ dt \le \int_0^x 2\cdot t^2 + t^2\cdot \sin(\tfrac{1}{t})\ dt \le \int_0^x 3\cdot t^2\ dt,$$

c'est-à-dire que

$$\frac{x^3}{3} \le \int_0^x 2\cdot t^2 + t^2\cdot \sin(\tfrac{1}{t})\ dt \le x^3.$$

Donc,

$$\lim_{x\to 0^+} \frac{x^2}{3} \le \lim_{x\to 0^+} \frac{f(x) - f(0)}{x - 0} \le \lim_{x\to 0^+} x^2.$$

D'où,

$$\lim_{x\to 0^+} \frac{f(x) - f(0)}{x - 0} = 0.$$

Par un mécanisme similaire, nous obtenons que

$$\lim_{x\to 0^-} \frac{f(x) - f(0)}{x - 0} = 0.$$

Finalement, la fonction f est dérivable sur $\mathbb{R}$ et

$$f'(x) = \begin{cases} 2\cdot x^2 + x^2\cdot \sin(\frac{1}{x}) & \text{si } x > 0 \\ 0 & \text{si } x = 0. \\ 2\cdot x^2 - x^2\cdot \sin(\frac{1}{x}) & \text{si } x < 0 \end{cases}$$

Il reste à montrer que cette fonction f' est localement Lipschitzienne sur $\mathbb{R}$.

Pour tout x non nul, la fonction f' est dérivable car elle est composée de fonctions dérivables. Ainsi, pour tout x non nul, f' est localement Lipschitzienne. Il nous reste à montrer que f' l'est également en $x = 0$. Pour cela, nous devons trouver des constantes $\varepsilon > 0$ et $K > 0$ telles que

$$| f'(x) - f'(y) | \le K\cdot | x - y |$$

pour tout x et y dans l'intervalle $]-\varepsilon,\varepsilon[$.

Si x et y sont simultanément positifs ou négatifs, la condition de Lipschitz est vérifiée grâce au caractère Lipschitzien des fonctions élémentaires. Prenons alors $x > 0$ et $y < 0$, d'où

$$| f'(x) - f'(y) | = | 2\cdot x^2 + x^2\cdot \sin(\tfrac{1}{x}) - 2\cdot y^2 + y^2\cdot \sin(\tfrac{1}{y}) |.$$

Par l'inégalité triangulaire, nous avons

$$| f'(x) - f'(y) | \le 2\cdot | x^2 - y^2 | + | x^2\cdot \sin(\tfrac{1}{x}) + y^2\cdot \sin(\tfrac{1}{y}) |.$$

Réécrivons le second terme.

$$| x^2\cdot \sin(\tfrac{1}{x}) + y^2\cdot \sin(\tfrac{1}{y}) | = | x^2\cdot \sin(\tfrac{1}{x}) - y^2\cdot \sin(-\tfrac{1}{y}) |.$$

Les nombres x et (-y) sont strictement positifs. Sans perte de généralité, nous supposerons que $0 < x < (-y) < 1$ car les points doivent être pris dans un voisinage de 0.

D'où, la fonction $(x^2 \cdot \sin(\frac{1}{x}))$ est dérivable sur $[x,(-y)]$. Le théorème des accroissements finis nous indique qu'il existe un point z de $[x,(-y)]$ tel que

$$\left| x^2 \cdot \sin(\tfrac{1}{x}) - y^2 \cdot \sin(-\tfrac{1}{y}) \right| = | x + y | \cdot \left| (z^2 \cdot \sin(\tfrac{1}{z}))' \right|$$
$$\leq | x - y | \cdot \left| 2 \cdot z \cdot \sin(\tfrac{1}{z}) - \cos(\tfrac{1}{z}) \right|$$
$$\leq 3 \cdot | x - y | .$$

La fonction $(\cdot)^2$ étant localement Lipschitzienne en 0, il existe des constantes $\varepsilon' > 0$ et $K' > 0$ telles que

$$| x^2 - y^2 | \leq K' \cdot | x - y |$$

pour tout x et y dans l'intervalle $]-\varepsilon', \varepsilon'[$.

En prenant $\varepsilon = \min(1, \varepsilon')$ et $K = 2 \cdot K' + 3$, nous obtenons que f est une fonction localement Lipschitzienne en 0.

En conclusion, la fonction f est de classe $C^{1,1}$ sur $\mathbb{R}$. Nous pouvons donc envisager la matrice Hessienne généralisée de f en un point de $\mathbb{R}$.

2°) Utilisons cette fonction f pour obtenir le contre-exemple voulu.

Remarquons que 0 est un point de minimum local pour la fonction f. En effet, sachant que $t^2 \leq \phi(t)$ pout tout $t \geq 0$, nous déduisons que $\frac{|x|^3}{3} \leq f(x)$ pour tout x dans $\mathbb{R}$. D'où $f(x) \geq 0$ pour tout x dans $\mathbb{R}$. Or, nous savons que $f(0) = 0$. Donc 0 est un point de minimum local de f.

Notons également que $\partial^2 f(0) = [-1,1]$.

Rappelons que $\Omega^c_{\nabla f} = \{ x \in \mathbb{R} \mid x \neq 0 \}$, et que, par définition,

$$\partial^2 f(0) = \mathrm{co}\,[\partial^2(f,0)],$$

avec $\partial^2(f,0) = \{ x \in \mathbb{R} \mid \exists (x_n)_{n=1}^{\infty} \subseteq \Omega^c_{\nabla f}$ qui converge vers 0 et telle que $x = \lim_{n\to\infty} f''(x_n) \}$.

Commençons par voir que $[-1,1]$ est inclus dans $\partial^2 f(0)$.

Prenons la suite $(x_n)_{n=1}^{\infty}$ définie pour tout $n \geq 1$ par

$$x_n = \frac{1}{2\cdot n\cdot \pi}.$$

Comme $x_n > 0$, nous calculons que

$$f''(x_n) = 4\cdot x_n + 2\cdot x_n\cdot \sin(\tfrac{1}{x_n}) - \cos(\tfrac{1}{x_n}).$$

C'est-à-dire que

$$\begin{aligned} f''(x_n) &= \frac{2}{n\cdot\pi} + \frac{1}{n\cdot\pi}\cdot \sin(2\cdot n\cdot\pi) - \cos(2\cdot n\cdot\pi) \\ &= \frac{2}{n\cdot\pi} - 1. \end{aligned}$$

D'où,

$$\lim_{n\to\infty} f''(x_n) = -1.$$

De la même façon, en prenant la suite $(x_n)_{n=1}^{\infty}$ définite par

$x_n = \frac{1}{(2n+1)\cdot\pi}$, nous obtenons que $\lim_{n\to\infty} f''(x_n) = 1$.

C'est-à-dire que les points 1 et (-1) appartiennent à $\partial^2(f,0)$.

D'où, $[-1,1]$ est contenu dans $\partial^2 f(0)$.

Montrons que $\partial^2 f(0)$ est inclus dans $[-1,1]$.

Soit $(x_n)_{n=1}^{\infty}$ une suite de points de $\Omega^c_{\nabla f}$ qui converge vers 0.

Soit $n \geq 1$. Si $x_n > 0$, alors

$$f''(x_n) = 4\cdot x_n + 2\cdot x_n\cdot \sin(\tfrac{1}{x_n}) - \cos(\tfrac{1}{x_n}).$$

Et si $x_n < 0$,

$$f''(x_n) = 4\cdot x_n + 2\cdot x_n\cdot \sin(\tfrac{1}{x_n}) + \cos(\tfrac{1}{x_n}).$$

En prenant la limite pour n tendant vers l'infini sur $f''(x_n)$, nous obtenons un nombre dont nous ne pouvons exprimer la valeur, mais nous savons que cette valeur appartient à l'intervalle $[-1,1]$.

D'où, l'inclusion $\partial^2 f(0) \subseteq [-1,1]$ est nécessairement vraie.

Finalement, nous avons que $\partial^2 f(0) = [-1,1]$.

Ainsi, bien que 0 soit un point de minimum local pour la fonction f, pour tout d dans $\mathbb{R}$, l'inégalité $d^T \cdot A \cdot d \geq 0$, n'est pas vérifiée quel que soit A dans $\partial^2 f(0)$ car $\partial^2 f(0) = [-1,1]$.

■

Commentaire 4.4.3 :

L'inégalité

$$\min \{ d^T \cdot A \cdot d \mid A \in \partial^2 f(x_o) \} \geq 0 \tag{4.10}$$

n'est pas toujours correcte quelque soit d dans $\mathbb{R}^p$.

Pour montrer cela, considérons les ensembles $\mathcal{A}(d)$ défini pour d appartenant à $\mathbb{R}^p$ par

$$\mathcal{A}(d) = \{ A \in \partial^2 f(x_o) \mid d^T \cdot A \cdot d \geq 0 \},$$

c'est-à-dire le plus grand ensemble de matrices de $\partial^2 f(x_o)$ telles que $d^T \cdot A \cdot d \geq 0$. Nous allons voir que cet ensemble dépend du point d choisi dans $\mathbb{R}^p$. Ainsi, l'inégalité (4.10) n'est pas toujours vérifiée quel que soit d dans $\mathbb{R}^p$ car $\mathcal{A}(d)$ ne recouvre par $\partial^2 f(x_o)$ tout entier. Utilisons un exemple pour montrer la dépendance de $\mathcal{A}(d)$.

Soit g une fonction g définie sur $\mathbb{R}^2$ et à valeurs réelles par

$$g(x,y) = f(x) + f(y)$$

pour tout (x,y) dans $\mathbb{R}^2$, où f est la fonction utilisée au commentaire 4.4.2.

Nous admettrons que cette fonction g est de classe $C^{1,1}$ sur $\mathbb{R}^2$ (cette propriété découle du caractère $C^{1,1}$ de la fonction f).

Remarquons que le point (0,0) est un point de minimum local de la fonction g car 0 est un point de minimum local de la fonction f.

Montrons que $\partial^2 g(0,0) = \left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mid a \text{ et } b \in [-1,1] \right\}$.

Par définition, $\partial^2 g(0,0) = \mathrm{co}\,[\partial^2(g,(0,0))]$

avec $\partial^2(g,(0,0)) = \{ A \in \mathcal{M}(2,2) \mid \exists (x_n,y_n)_{n=1}^{\infty} \subseteq \Omega^c_{\nabla g}$ convergente vers (0,0) telle que $A = \lim_{n\to\infty} \nabla^2 g(x_n,y_n) \}$.

De plus, par définition de la fonction g,

$$\Omega^c_{\nabla g} = \{ (x,y) \mid (x,y) \neq (0,0) \} = \Omega^c_{\nabla f} \times \Omega^c_{\nabla f}.$$

Commençons par voir que

$$\partial^2 g(0,0) \text{ est inclus dans } \left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mid a \text{ et } b \in \partial^2 f(0) \right\}.$$

Soit $(x_n,y_n)_{n=1}^{\infty}$ une suite de points de $\Omega^c_{\nabla g}$ qui converge vers (0,0). Pour tout $n \geq 1$, nous avons

$$\nabla^2 g(x_n,y_n) = \begin{bmatrix} \dfrac{\partial^2 g}{\partial x^2}(x_n,y_n) & \dfrac{\partial^2 g}{\partial x \partial y}(x_n,y_n) \\ \dfrac{\partial^2 g}{\partial y \partial x}(x_n,y_n) & \dfrac{\partial^2 g}{\partial y^2}(x_n,y_n) \end{bmatrix}.$$

Nous calculons que

$$\frac{\partial^2 g}{\partial x^2}(x_n,y_n) = \frac{\partial^2 (f(x_n)+g(y_n))}{\partial x^2} = f''(x_n),$$

$$\frac{\partial^2 g}{\partial x \partial y}(x_n,y_n) = \frac{\partial^2 (f(x_n)+g(y_n))}{\partial x \partial y} = 0, \text{ et}$$

$$\frac{\partial^2 g}{\partial y^2}(x_n,y_n) = \frac{\partial^2 (f(x_n)+g(y_n))}{\partial y^2} = f''(y_n).$$

D'où,

$$\nabla^2 g(x_n,y_n) = \begin{bmatrix} f''(x_n) & 0 \\ 0 & f''(y_n) \end{bmatrix}.$$

Ainsi, par passage à la limite, nous obtenons que

$$\lim_{n\to\infty} \nabla^2 g(x_n,y_n) = \begin{bmatrix} a & o \\ o & b \end{bmatrix},$$

où a et b sont des éléments de $\partial^2(f,0)$.

Autrement dit, suite à la caractérisation de l'enveloppe convexe d'un ensemble,

$$\partial^2 g(0,0) \subseteq \left\{ \begin{bmatrix} a & o \\ o & b \end{bmatrix} \mid a \text{ et } b \in \partial^2 f(0) \right\}.$$

Montrons que $\left\{ \begin{bmatrix} a & o \\ o & b \end{bmatrix} \mid a \text{ et } b \in \partial^2 f(0) \right\}$ est inclus dans $\partial^2 g(0,0)$.

Par la caractérisation de l'enveloppe convexe d'un ensemble, il suffit de montrer que

$$\left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mid a \text{ et } b \in \partial^2(f,0) \right\} \text{ est inclus dans } \partial^2(g,(0,0)).$$

Soient a et b appartenant à $\partial^2(f,0)$. Ainsi, il existe des suites $(x_n)_{n=1}^{\infty}$ et $(y_n)_{n=1}^{\infty}$ de nombres réels non nuls qui convergent vers 0 telles que $a = \lim_{n\to\infty} f''(x_n)$ et $b = \lim_{n\to\infty} f''(y_n)$.

D'où,

$$\lim_{n\to\infty} \begin{bmatrix} f''(x_n) & 0 \\ 0 & f''(y_n) \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}.$$

Autrement dit

$$\lim_{n\to\infty} \nabla^2 g(x_n,y_n) = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}.$$

Comme la suite $(x_n,y_n)_{n=1}^{\infty}$ appartient à $\Omega_{\nabla g}^{c}$ et qu'elle converge vers $(0,0)$, la matrice $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ appartient à $\partial^2(g,(0,0))$.

Donc, $\partial^2 g(0,0) = \left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mid a \text{ et } b \in \partial^2 f(0) \right\}$.

Or, $\partial^2 f(0) = [-1,1]$, d'où, nous obtenons finalement que

$$\partial^2 g(0,0) \subseteq \left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mid a \text{ et } b \in [-1,1] \right\}.$$

Cette expression de $\partial^2 g(0,0)$ entraîne que

$$\mathscr{A}(d) = \left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mid a \text{ et } b \in [-1,1] \text{ et } (a\cdot d_1^2 + b\cdot d_2^2) \geq 0 \right\}$$

pour tout $d = (d_1,d_2)^T$ appartenant à $\mathbb{R}^2$.

Si nous prenons $d^* = (1,0)^T$ et $d^{**} = (0,1)^T$, nous remarquons que $\mathcal{A}(d^*)$ est différent de $\mathcal{A}(d^{**})$.

■

4.4.B: Le cas des problèmes avec contraintes.

Soit f une fonction définie sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles. Soient g_i et h_j des fonctions définies sur $\mathbb{R}^p$ et à valeurs réelles quels que soient i dans un ensemble fini de nombres naturels I et j dans un ensemble fini de nombres naturels E. Notons par m, respectivement par n, le nombre d'éléments de I, respectivement de E.

Supposons que chacune de ces fonctions soit de classe $\mathcal{C}^{1,1}$ sur son domaine de définition.

La forme générale du problème avec contraintes est

$$(PAC)\quad \begin{cases} \text{Minimiser } f(x) \\ \text{sous conditions: } g_i(x) \le 0 \text{ pour tout } i \text{ dans } I \\ \qquad\qquad\qquad\quad h_j(x) = 0 \text{ pour tout } j \text{ dans } E \\ \qquad\qquad\qquad\quad x \text{ appartient à } \mathcal{O}. \end{cases}$$

Soit x_o un point de minimum local de ce problème (PAC).

Avant de développer des conditions nécessaires d'optimalité du second ordre pour ce problème (PAC), nous allons rappeler des conditions nécessaires du premier ordre que nous connaissons (plus de détails sont exposés dans le paragraphe 1.8).

Comme la fonction f et que chaque fonction g_i et h_j sont continûment différentiables sur leur domaine, nous savons que la *condition nécessaire d'optimalité de Fritz-John* est correcte.

Autrement dit, il existe un vecteur $(\lambda_0,\lambda_1,\dots,\lambda_m,\mu_1,\dots,\mu_n)$ de $\mathbb{R}^{m+n+1}$ non nul tel que

$$(FJ)\quad \begin{cases} \lambda_0\cdot\nabla f(x_0) + \sum\limits_{i\in I} \lambda_i\cdot\nabla g_i(x_0) + \sum\limits_{j\in E} \mu_j\cdot\nabla h_j(x_0) = 0, \\ \lambda_0 \geq 0, \text{ et pour tout } i \text{ dans } I, \\ \lambda_i \geq 0 \text{ et } \lambda_i\cdot g_i(x_0) = 0. \end{cases}$$

De plus, sous une *condition de qualification des contraintes* (que nous noterons de manière générale par *(QC)*), la condition de Fritz-John conduit à la *condition nécessaire d'optimalité de Kuhn-Tucker*.

C'est-à-dire qu'il existe un vecteur $(\lambda_1,\dots,\lambda_m,\mu_1,\dots,\mu_n)$ de $\mathbb{R}^{m+n}$ tel que

$$(KT)\quad \begin{cases} \nabla f(x_0) + \sum\limits_{i\in I} \lambda_i\cdot\nabla g_i(x_0) + \sum\limits_{j\in E} \mu_j\cdot\nabla h_j(x_0) = 0, \\ \lambda_0 \geq 0, \text{ et pour tout } i \text{ dans } I, \\ \lambda_i \geq 0 \text{ et } \lambda_i\cdot g_i(x_0) = 0. \end{cases}$$

Considérons un vecteur $(\lambda,\mu) = (\lambda_1,\dots,\lambda_m,\mu_1,\dots,\mu_n)$ de Kuhn-Tucker.

Nous aurons également besoin de l'ensemble $G(\lambda)$ défini par

$$G(\lambda) = \left\{ x \in \mathbb{R}^p \;\middle|\; \begin{array}{l} g_i(x) = 0 \text{ pour tout } i \text{ de } I \text{ tel que } \lambda_i > 0 \\ g_i(x) \leq 0 \text{ pour tout } i \text{ de } I \text{ tel que } \lambda_i = 0 \\ h_j(x) = 0 \text{ pour tout } j \text{ de } E \end{array} \right\}$$

Ainsi, nous pouvons envisager $\mathcal{T}_\lambda(x_o)$ le cône tangent de Bouligand de $\mathcal{G}(\lambda)$ au point x_o. D'où, $\mathcal{T}_\lambda(x_o)$ est l'ensemble des points d de $\mathbb{R}^p$ pour lesquels il existe une suite $(d_k)_{k=1}^{\infty}$ dans $\mathbb{R}^p$ qui converge vers d, et il existe une suite $(\alpha_k)_{k=1}^{\infty}$ de nombres réels qui converge vers 0 par valeurs positives telles que $(x_o + \alpha_k \cdot d_k)$ appartient à $\mathcal{G}(\lambda)$ quel que soit $k \geq 1$.

Autrement dit,

$$\mathcal{T}_\lambda(x_o) = \left\{ d \in \mathbb{R}^p \;\middle|\; \begin{array}{l} \exists (d_k)_{k=1}^{\infty} \subseteq \mathbb{R}^p \text{ convergente vers } d \\ \exists (\alpha_k)_{k=1}^{\infty} \subseteq \mathbb{R} \text{ convergente vers 0 par} \\ \qquad\qquad \text{valeurs positives} \\ \text{telles que } (x_o + \alpha_k \cdot d_k) \in \mathcal{G}(\lambda) \ \forall k \geq 1 \end{array} \right\}.$$

Rappelons aussi le Lagrangien $L(x,\lambda,\mu)$ du problème *(PAC)*, c'est-à-dire

$$L(x,\lambda,\mu) = f(x) + \sum_{i \in I} \lambda_i \cdot g_i(x) + \sum_{j \in E} \mu_j \cdot h_j(x).$$

Les propriétés 4.1.2 et 4.1.3 entraînent que la fonction $L(\cdot,\lambda,\mu)$ est de classe $\mathcal{C}^{1,1}$ sur $\mathcal{O}$. D'où, nous pouvons prendre la matrice Hessienne généralisée $\partial^2_{xx} L(x_o,\lambda,\mu)$ de la fonction $L(\cdot,\lambda,\mu)$ au point x_o.

Ces différentes remarques nous permettent d'établir le théorème suivant qui fournit des conditions nécessaires du second ordre pour le problème *(PAC)*.

Théorème 4.4.4 : Soit x_o appartenant à $\mathcal{O}$,

si x_o est un point de minimum local pour le problème *(PAC)* et si la condition *(QC)* de qualification de contraintes est vérifiée au point x_o,

alors pour tout vecteur (λ,μ) de Kuhn-Tucker et pour tout point d de $\mathcal{T}_\lambda(x_o)$, il existe une matrice A de $\partial^2_{xx}L(x_o,\lambda,\mu)$ telle que $d^T\cdot A\cdot d \geq 0$.

Preuve :

Soit $(\lambda,\mu) = (\lambda_1,\ldots,\lambda_m,\mu_1,\ldots,\mu_n)$ un vecteur de Kuhn-Tucker et soit d un point de $\mathcal{T}_\lambda(x_o)$.

Par définition de $\mathcal{T}_\lambda(x_o)$, il existe une suite $(d_k)_{k=1}^{\infty}$ dans $\mathbb{R}^p$ qui converge vers d, et il existe une suite $(\alpha_k)_{k=1}^{\infty}$ de nombres réels qui converge vers 0 par valeurs positives telles que $(x_o + \alpha_k\cdot d_k)$ appartient à $G(\lambda)$ quel que soit $k \geq 1$.

Soit $k \geq 1$. Appliquons le théorème 4.3.3 à la fonction $L(\cdot,\lambda,\mu)$ sur l'intervalle $[x_o,x_o+\alpha_k\cdot d_k]$.

Il indique qu'il existe un nombre α_k^* compris entre 0 et α_k, et qu'il existe une matrice A_k appartenant à $\partial^2_{xx}L(x_o+\alpha_k\cdot d_k,\lambda,\mu)$ tels que

$$L(x_o+\alpha_k\cdot d_k,\lambda,\mu) = L(x_o,\lambda,\mu) + \langle \nabla_x L(x_o,\lambda,\mu) , \alpha_k\cdot d_k \rangle + \frac{1}{2}\cdot\alpha_k^2 \cdot d_k^T\cdot A_k\cdot d_k. \quad (4.11)$$

Or, x_o est un point de minimum du problème *(PAC)*. Ainsi, grâce à la condition nécessaire du premier ordre *(KT)*, nous avons

$$L(x_o,\lambda,\mu) = f(x_o), \text{ et}$$

$$\nabla_x L(x_o,\lambda,\mu) = 0.$$

De plus, par définition de *G(λ)*, nous avons

$$L(x_o + \alpha_k \cdot d_k,\lambda,\mu) = f(x_o + \alpha_k \cdot d_k).$$

D'où, suite à ces égalités, (4.11) devient

$$f(x_o + \alpha_k \cdot d_k) = f(x_o) + \frac{1}{2} \cdot \alpha_k^2 \cdot d_k^T \cdot A_k \cdot d_k. \tag{4.12}$$

Comme x_o est un point de minimum local du problème *(PAC)*,

$$f(x_o + \alpha_k \cdot d_k) \geq f(x_o).$$

D'où, (4.12) entraîne que

$$d_k^T \cdot A_k \cdot d_k \geq 0.$$

Finalement, en reprenant le raisonnement du théorème 4.4.1, il existe une sous-suite $(A_{q_k})_{k=1}^{\infty}$ extraite de $(A_k)_{k=1}^{\infty}$ qui converge vers une matrice A de $\mathcal{M}(p,p)$ telle que A appartient à $\partial^2_{xx} L(x_o,\lambda,\mu)$ et $d^T \cdot A \cdot d \geq 0$.

□

Si la condition *(QC)* de qualification des contraintes n'est pas satisfaite, la condition *(KT)* de Kuhn-Tucker n'est pas nécessairement vérifiée.

Mais, la condition de Fritz-John est correcte. Ainsi, il existe un vecteur non nul $(\lambda_o,\lambda,\mu) = (\lambda_o,\lambda_1,\ldots,\lambda_m,\mu_1,\ldots,\mu_n)$ de $\mathbb{R}^{m+n+1}$ vérifiant les conditions *(FJ)*.

Avec le Lagrangien $L(x,\lambda_o,\lambda,\mu)$ défini par

$$L(x,\lambda_o,\lambda,\mu) = \lambda_o \cdot f(x) + \sum_{i \in I} \lambda_i \cdot g_i(x) + \sum_{j \in E} \mu_j \cdot h_j(x),$$

des conditions nécessaires du second ordre sont fournies par le théorème suivant.

Théorème 4.4.5 : Soit x_o appartenant à $\mathcal{O}$,

si x_o est un point de minimum local pour le problème *(PAC)*

alors pour tout vecteur (λ_o,λ,μ) de Fritz-John et pour tout point d de $\mathcal{T}_\lambda(x_o)$, il existe une matrice A de $\partial^2_{xx}L(x_o,\lambda_o,\lambda,\mu)$ telle que $d^T \cdot A \cdot d \geq 0$.

Nous n'en développerons pas la démonstration car elle est identique à celle du théorème 4.4.4.

A présent, nous allons essayer de formuler le théorème 4.4.4 de manière différente. Supposons alors que les conditions *(QC)* de qualifications de contraintes soient vérifiées. Commençons par montrer la propriété qui suit.

Propriété 4.4.6 :

$$\mathcal{T}_\lambda(x_o) \subseteq \left\{ d \in \mathbb{R}^p \;\middle|\; \begin{array}{l} \langle \nabla g_i(x_o),d\rangle = 0 \text{ pour tout } i \text{ tel que } \lambda_i > 0 \\ \langle \nabla g_i(x_o),d\rangle \leq 0 \text{ pour tout } i \text{ tel que } \lambda_i = 0 \\ \qquad\qquad\qquad\qquad \text{et } g_i(x_o) = 0 \\ \langle \nabla h_j(x_o),d\rangle = 0 \text{ pour tout } j \end{array} \right\}.$$

Preuve :

Soit d appartenant à $T_\lambda(x_o)$.

Nous devons montrer que

$\langle \nabla g_i(x_o), d \rangle = 0$ pour tout i de I tel que $\lambda_i > 0$,

$\langle \nabla g_i(x_o), d \rangle \leq 0$ pour tout i de I tel que $\lambda_i = 0$ et $g_i(x_o) = 0$,

et

$\langle \nabla h_j(x_o), d \rangle = 0$ pour tout j de E.

Si d = 0, ces égalités et cette inégalité sont trivialement vraies. Supposons donc que d est non nul.

Comme les fonctions g_i et h_j sont continûment différentiables, nous avons

$$\lim_{x \to x_o} \frac{g_i(x) - g_i(x_o)}{\| x - x_c \|} = \lim_{x \to x_o} \frac{\langle \nabla g_i(x_o), d \rangle}{\| x - x_o \|},$$

pour tout i dans I; et

$$\lim_{x \to x_o} \frac{h_j(x) - h_j(x_o)}{\| x - x_c \|} = \lim_{x \to x_o} \frac{\langle \nabla h_j(x_o), d \rangle}{\| x - x_o \|},$$

pour tout j dans E.

Par définition de $T_\lambda(x_o)$, il existe une suite $(d_k)_{k=1}^{\infty}$ dans $\mathbb{R}^p$ qui converge vers d, et il existe une suite $(\alpha_k)_{k=1}^{\infty}$ de nombres réels qui converge vers 0 par valeurs positives telles que $(x_o + \alpha_k \cdot d_k)$ appartient à $G(\lambda)$ quel que soit $k \geq 1$.

La suite $(x_o + \alpha_k \cdot d_k)_{n=1}^{\infty}$ ainsi obtenue converge évidemment vers x_o.

D'où,

$$\lim_{k\to\infty} \frac{g_i(x_o + \alpha_k \cdot d_k) - g_i(x_o)}{\alpha_k \cdot \|d_k\|} = \lim_{k\to\infty} \frac{\langle \nabla g_i(x_o), d_k\rangle}{\|d_k\|}$$

$$= \frac{\langle \nabla g_i(x_o), d\rangle}{\|d\|},$$

pour tout i dans I; et

$$\lim_{k\to\infty} \frac{h_j(x_o + \alpha_k \cdot d_k) - h_j(x_o)}{\alpha_k \cdot \|d_k\|} = \lim_{k\to\infty} \frac{\langle \nabla h_j(x_o), d_k\rangle}{\|d_k\|}$$

$$= \frac{\langle \nabla h_j(x_o), d\rangle}{\|d\|},$$

pour tout j dans E.

Soit i dans I.

Vérifions que $\langle \nabla g_i(x_o), d\rangle = 0$ si $\lambda_i > 0$.

Comme $(x_o + \alpha_k \cdot d_k)$ appartient à $G(\lambda)$ quel que soit $k \geq 1$, nous savons que $g_i(x_o + \alpha_k \cdot d_k) = 0$.

De plus, les condition *(KT)* entraînent que $\lambda_i \cdot g_i(x_o) = 0$. D'où, $g_i(x_o) = 0$ car $\lambda_i > 0$.

Ainsi,

$$\lim_{k\to\infty} \frac{g_i(x_o + \alpha_k \cdot d_k) - g_i(x_o)}{\alpha_k \cdot \|d_k\|} = 0.$$

Autrement dit,

$$\langle \nabla g_i(x_o), d\rangle = 0$$

car d est non nul.

Vérifions que $\langle \nabla g_i(x_o),d\rangle \leq 0$ pour tout i tel que $\lambda_i = 0$ et $g_i(x_o) = 0$.

Comme $g_i(x_o) = 0$,

$$\lim_{k\to\infty} \frac{g_i(x_o + \alpha_k \cdot d_k) - g_i(x_o)}{\alpha_k \cdot \|d_k\|} = \lim_{k\to\infty} \frac{g_i(x_o + \alpha_k \cdot d_k)}{\alpha_k \cdot \|d_k\|}.$$

Or, $(x_o + \alpha_k \cdot d_k)$ appartienant à $G(\lambda)$ quel que soit $k \geq 1$, il suit que $g_i(x_o + \alpha_k \cdot d_k) \leq 0$. Et comme $\alpha_k > 0$ et que $\|d_k\| > 0$, nous avons

$$\lim_{k\to\infty} \frac{g_i(x_o + \alpha_k \cdot d_k)}{\alpha_k \cdot \|d_k\|} \leq 0.$$

Autrement dit,

$$\langle \nabla g_i(x_o),d\rangle \leq 0$$

car d est non nul.

Soit j dans E.

Vérifions que $\langle \nabla h_j(x_o),d\rangle = 0$ pour tout j.

Nous savons que $h_j(x_o) = 0$.

De plus, $h_j(x_o + \alpha_k \cdot d_k) = 0$ car $(x_o + \alpha_k \cdot d_k)$ appartient à $G(\lambda)$.

D'où,

$$\lim_{k\to\infty} \frac{h_j(x_o + \alpha_k \cdot d_k) - h_j(x_o)}{\alpha_k \cdot \|d_k\|} = 0.$$

Autrement dit,

$$\langle \nabla h_j(x_o),d\rangle = 0$$

car d est non nul.

□

L'inclusion inverse de la propriété 4.4.6 n'est pas toujours vérifiée. Pour la rendre exacte, nous allons devoir ajouter des conditions de régularité du second-ordre, que nous noterons par *(CR)*. Un exemple de conditions *(CR)* est fourni par la propriété qui suit.

Propriété 4.4.7: Si les vecteurs $\nabla g_i(x_o)$ et $\nabla h_j(x_o)$ sont linéairement indépendants quels que soient i dans $I(x_o)$ et j dans E,

où $I(x_o) = \{\ i \in I \mid g_i(x_o) = 0\ \}$,

alors

$$T_\lambda(x_o) = \left\{ d \in \mathbb{R}^p \;\middle|\; \begin{array}{l} \langle \nabla g_i(x_o), d\rangle = 0 \text{ pour tout } i \text{ tel que } \lambda_i > 0 \\ \langle \nabla g_i(x_o), d\rangle \leq 0 \text{ pour tout } i \text{ tel que } \lambda_i = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{et } g_i(x_o) = 0 \\ \langle \nabla h_j(x_o), d\rangle = 0 \text{ pour tout } j \end{array} \right\}.$$

Preuve :

Il suffit de montrer que

$$\left\{ d \in \mathbb{R}^p \;\middle|\; \begin{array}{l} \langle \nabla g_i(x_o), d\rangle = 0 \text{ pour tout } i \text{ tel que } \lambda_i > 0 \\ \langle \nabla g_i(x_o), d\rangle \leq 0 \text{ pour tout } i \text{ tel que } \lambda_i = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{et } g_i(x_o) = 0 \\ \langle \nabla h_j(x_o), d\rangle = 0 \text{ pour tout } j \end{array} \right\} \subseteq T_\lambda(x_o)$$

car l'inclusion inverse est vraie en vertu de la propriété 4.4.6.

Soit d appartenant à $\mathbb{R}^p$ tel que

$\langle \nabla g_i(x_o), d\rangle = 0$ pour tout i tel que $\lambda_i > 0$

$\langle \nabla g_i(x_o), d\rangle \leq 0$ pour tout i tel que $\lambda_i = 0$ et $g_i(x_o) = 0$.

$\langle \nabla h_j(x_o), d\rangle = 0$ pour tout j

Posons $R_a^{p+1} = \{ (x,\theta) \in \mathbb{R}^{p+1} \mid$ x est admissible pour *(PAC)* $\}$.

Notons $E \cup I(x_o)$ par $\{1,\dots,r\}$. Remarquons que $r \leq p$, car si ce n'est pas le cas, les vecteurs $\nabla g_i(x_o)$ et $\nabla h_j(x_o)$ ne sont pas linéairement indépendants quels que soient i dans $I(x_o)$ et j dans E.

Considérons la fonction auxiliaire Φ définie sur R_a^{p+1} et à valeurs dans $\mathbb{R}^p$ par

$$\Phi_i(x,\theta) = \begin{cases} h_i(x) - \theta\cdot\langle \nabla h_i(x_o), d\rangle & \text{si i est dans E} \\ g_i(x) - \theta\cdot\langle \nabla g_i(x_o), d\rangle & \text{si i est dans } I(x_o) \\ \langle x - x_o, b_i\rangle - \theta\cdot\langle b_i, d\rangle & \text{si i est dans } \{r+1,\dots,p\} \end{cases},$$

avec (x,θ) appartenant à R_a^{p+1}.

Les nombres $b_{r+1},\dots,b_r$ seront pris d'une certaine manière plus loin. Pour l'instant, considérons les comme des nombres quelconques de $\mathbb{R}^p$. Ils nous permettent d'obtenir un système de p équations à (p+1) inconnues.

Schématiquement, nous noterons ce système par

$$\Phi(x,\theta) = 0. \tag{4.13}$$

Remarquons que le point $(x_o,0)$ de R_a^{p+1} est solution de ce système. La matrice Jacobienne de cette fonction continûment différentiable par rapport à la variable x est

$$J_x\Phi(x,\theta) = \begin{bmatrix} [\nabla h_1(x)]^T \\ \vdots \\ [\nabla h_n(x)]^T \\ [\nabla g_1(x)]^T \\ \vdots \\ [\nabla g_{r-n}(x)]^T \\ b_{r+1} \\ \vdots \\ b_p \end{bmatrix}.$$

Comme les vecteurs $\nabla g_i(x_o)$ et $\nabla h_j(x_o)$ sont linéairement indépendants quels que soient i dans $I(x_o)$ et j dans E, nous pouvons choisir les vecteurs b_{r+1}, ... , b_p tels que la matrice $[J_x\Phi(x_o,0)]^T$ est inversible.

Ainsi, par le théorème des fonctions implictes, il existe une fonction continûment différentiable $x(\cdot)$ définie sur un voisinage réel V_θ de 0 et à valeurs dans un voisinage V_x de x_o telle que $(x(\theta),\theta)$ est solution du système (4.13).

De plus, comme $\Phi(x(\theta),\theta) = 0$ sur le voisinage V_θ de 0, nous déduisons que, pour tout i compris entre 1 et p,

$$\frac{d\Phi_i}{d\theta}(x(\theta),\theta) = 0. \tag{4.14}$$

Ainsi, par les règles de dérivation,

$$\sum_{j=1}^{p} \left(\frac{\partial\Phi_i}{\partial x_j}(x(\theta),\theta)\cdot\frac{\partial x_j}{\partial\theta} + \frac{\partial\Phi_i}{\partial\theta}(x(\theta),\theta) \right) = 0.$$

Or, si i appartient à E,

$$\frac{\partial \Phi_i}{\partial x_j}(x(\theta),\theta) = \frac{\partial h_i}{\partial x_j}(x(\theta)) \text{ et } \frac{\partial \Phi_i}{\partial \theta}(x(\theta),\theta) = -\langle \nabla h_i(x_o),d\rangle,$$

si i appartient à $I(x_o)$,

$$\frac{\partial \Phi_i}{\partial x_j}(x(\theta),\theta) = \frac{\partial g_i}{\partial x_j}(x(\theta)) \text{ et } \frac{\partial \Phi_i}{\partial \theta}(x(\theta),\theta) = -\langle \nabla g_i(x_o),d\rangle,$$

et enfin, si i appartient à $\{r+1,\ldots,p\}$,

$$\frac{\partial \Phi_i}{\partial x_j}(x(\theta),\theta) = (b_i)_j \text{ et } \frac{\partial \Phi_i}{\partial \theta}(x(\theta),\theta) = -\langle b_i,d\rangle.$$

D'où, matriciellement, (4.14) devient

$$[J_x\Phi(x(\theta),\theta)]^T\cdot\frac{dx}{d\theta} - [J_x\Phi(x_o,0)]^T\cdot d = 0.$$

Ainsi, si $\theta = 0$, nous avons

$$\frac{dx}{d\theta}\bigg|_{\theta=0} = d,$$

car $[J_x\Phi(x_o,0)]^T$ est une matrice inversible.

Autrement dit,

$$\lim_{\theta\to 0}\frac{x(\theta) - x_o}{\theta} = d.$$

A présent, considérons une suite $(\theta_k)_{k=1}^{\infty}$ de points réels qui vers 0 par valeurs positives. Comme la fonction $x(\cdot)$ est continue, la suite $(x(\theta_k))_{k=1}^{\infty}$ converge vers x_o. Ainsi, nous avons évidemment que

$$\lim_{k\to 0}\frac{x(\theta_k) - x_o}{\theta_k} = d.$$

Et de plus, il existe un nombre N de $\mathbb{N}_o$ tel que $x(\theta_k)$ appartient à $G(\lambda)$ pour tout $k \geq N$.

En effet, par définition de la fonction $x(\cdot)$, il existe un nombre N de $\mathbb{N}_o$ tel que θ_k appartient à V_θ pour tout $k \geq N$.

D'où,

$$\Phi(x(\theta_k),\theta_k) = 0$$

pour tout $k \geq N$.

Ainsi, pour tout i appartenant à $\{1,\ldots,p\}$, nous avons

$$\Phi_i(x(\theta_k),\theta_k) = 0,$$

quel que soit $k \geq N$.

D'où, si i appartient à E,

$$h_i(x(\theta_k)) - \theta_k \cdot \langle \nabla h_i(x_o),d \rangle = 0.$$

Or, d est tel que $\langle \nabla h_i(x_o),d \rangle = 0$, donc $h_i(x(\theta_k)) = 0$.

Et, si i appartient à $I(x_o)$,

$$g_i(x(\theta_k)) - \theta_k \cdot \langle \nabla g_i(x_o),d \rangle = 0.$$

Si de plus i est tel que $\lambda_i > 0$, nous savons que d est tel que $\langle \nabla g_i(x_o),d \rangle = 0$, donc $g_i(x(\theta_k)) = 0$.

Tandis que si i est tel que $\lambda_i = 0$ et $g_i(x_o) = 0$, alors $\langle \nabla g_i(x_o),d \rangle \leq 0$. Comme $\theta_k \geq 0$, nous obtenons que $\theta_k \cdot \langle \nabla g_i(x_o),d \rangle \leq 0$. Ainsi $g_i(x(\theta_k)) \leq 0$.

Sans perte de généralité, nous supposerons que $x(\theta_k)$ appartient à $G(\lambda)$ pour tout $k \geq 1$, sinon il suffit de prendre la suite $(\theta_k)_{k=1}^{\infty}$ à partir du rang N.

Prenons la suite $(d_k)_{k=1}^{\infty}$ de $\mathbb{R}^p$ définie pour tout $k \geq 1$ par

$$d_k = \frac{x(\theta_k) - x_o}{\theta_k}.$$

Ainsi, comme nous calculons que $(x_o + \theta_k \cdot d_k) = x(\theta_k)$, il suit que $(x_o + \theta_k \cdot d_k)$ appartient à $\mathcal{G}(\lambda)$.

La suite $(d_k)_{k=1}^{\infty}$ étant convergente vers d et la suite $(\theta_k)_{k=1}^{\infty}$ étant convergente vers 0 par valeurs positives, nous obtenons que d appartient à $\mathcal{T}_\lambda(x_o)$.

□

Cette propriété 4.4.7 nous permet de reformuler le théorème 4.4.4 de la manière suivante.

Théorème 4.4.8 : Soit x_o appartenant à $\mathcal{O}$,

si x_o est un point de minimum local pour le problème *(PAC)* et si les conditions *(QC)* de qualifications de contraintes et si les conditions *(CR)* sont vérifiées,

alors pour tout vecteur (λ,μ) de Kuhn-Tucker et pour tout point d de $\mathbb{R}^p$ tel que,

$\langle\nabla g_i(x_o),d\rangle = 0$ pour tout i tel que $\lambda_i > 0$,

$\langle\nabla g_i(x_o),d\rangle \leq 0$ pour tout i tel que $\lambda_i = 0$ et $g_i(x_o) = 0$,

$\langle\nabla h_j(x_o),d\rangle = 0$ pour tout j,

il existe une matrice A de $\partial^2_{xx}L(x_o,\lambda,\mu)$ telle que $d^T \cdot A \cdot d \geq 0$.

Finalement, nous obtenons des propriétés qui nous fournissent des renseignements similaires à celles pour le cas deux fois continûment différentiable.

Mais, chacune de ces propriétés ne nous fournissent des caractéristiques concernant l'une des matrices de la matrice Hessienne généralisée.

ANNEXES

Nous allons décrire les propriétés utilisées dans les chapitres précédents.

A) PREMIÈRE PARTIE.

Soit F une fonction définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}^q$. Soit v appartenant à $\mathbb{R}^q$. Considérons la fonction F_v définie sur un ouvert $\mathcal{O}$ de $\mathbb{R}^p$ et à valeurs réelles par $F_v(x) = \langle F(x),v\rangle$ pout tout x appartenant à $\mathcal{O}$. Soit x_o dans $\mathcal{O}$, nous avons les propriétés suivantes.

Propriété A.1 : Si F est différentiable en x_o,
alors F_v est différentiable en x_o,
et $\nabla F_v(x_o) = [\,JF(x_o)]^T \cdot v$.

Preuve :

Tout sera démontré si nous obtenons que

$$\lim_{x \to x_o} \frac{F_v(x) - F_v(x_o) - ([JF(x_o)]^T \cdot v) \cdot (x - x_o)}{\|x - x_o\|} = 0.$$

Réécrivons l'expression sur laquelle est calculée la limite.

$$\frac{F_v(x) - F_v(x_o) - ([JF(x_o)]^T \cdot v) \cdot (x - x_o)}{\|x - x_o\|}$$

$$= \frac{F(x)^T \cdot v - F(x_o)^T \cdot v - ([JF(x_o)]^T \cdot v) \cdot (x - x_o)}{\|x - x_o\|}$$

$$= \frac{\langle F(x) - F(x_o) - JF(x_o) \cdot (x - x_o) \, , \, v \rangle}{\|x - x_o\|}.$$

L'inégalité de Cauchy-Schwarz indique que

$$\frac{-\| F(x) - F(x_o) - JF(x_o) \cdot (x - x_o) \| \cdot \| v \|}{\|x - x_o\|}$$

$$\leq \frac{\langle F(x) - F(x_o) - JF(x_o) \cdot (x - x_o) \, , \, v \rangle}{\|x - x_o\|}$$

$$\leq \frac{\| F(x) - F(x_o) - JF(x_o) \cdot (x - x_o) \| \cdot \| v \|}{\|x - x_o\|}.$$

Comme F est différentiable en x_o, nous savons que

$$\lim_{x \to x_o} \frac{F(x) - F(x_o) - JF(x_o) \cdot (x - x_o)}{\|x - x_o\|} = 0.$$

D'où,

$$\lim_{x \to x_o} \frac{\| F(x) - F(x_o) - JF(x_o) \cdot (x - x_o) \|}{\|x - x_o\|} = 0.$$

Ainsi,

$$\lim_{x \to x_o} \frac{\langle\, F(x) - F(x_o) - JF(x_o)\cdot(x - x_o)\ ,\ v \,\rangle}{\|x - x_o\|} = 0.$$

Donc,

$$\lim_{x \to x_o} \frac{F_v(x) - F_v(x_o) - ([\,JF(x_o)]^T\cdot v)\cdot(x - x_o)}{\|x - x_o\|} = 0.$$

□

Désignons par Ω_F l'ensemble des points de $\mathcal{O}$ où F n'est pas différentiable, et par Ω_{F_v} l'ensemble des points de $\mathcal{O}$ où F_v n'est pas différentiable.

Grâce à la propriété A.1, nous obtenons directement le corollaire suivant.

Corollaire A.2 : Ω_{F_v} est inclus dans Ω_F.

Propriété A.3 : Si la fonction F est localement Lipschitzienne en x_o,
alors F_v est une fonction localement
Lipschitzienne en x_o.

Preuve :

Il faut trouver des constantes $\varepsilon_v > 0$ et $K_v > 0$ telles que

$|F_v(x) - F_v(y)| \le K_v\cdot\|x - y\|$ pour tout x et y dans $B(x_o,\varepsilon_v)$.

Par définition de F_v et par linéarité du produit scalaire par rapport à la première variable, il suit que

$$| F_v(x) - F_v(y) | = |\langle F(x) - F(y) , v\rangle| .$$

Ainsi, par l'inégalité de Cauchy-Schwarz, nous obtenons que

$$| F_v(x) - F_v(y) | \le \| F(x) - F(y) \| \cdot \| v \| .$$

Comme F est localement Lipschitzienne en x_o, nous savons qu'il existe des constantes $\varepsilon > 0$ et $K > 0$ telles que

$$\| F(x) - F(y) \| \le K \cdot \| x - y \| \text{ pour tout } x \text{ et } y \text{ dans } B(x_o,\varepsilon).$$

D'où, si x et y appartiennent à $B(x_o,\varepsilon)$, il suit que

$$| F_v(x) - F_v(y) | \le K \cdot \| v \| \cdot \| x - y \| .$$

Donc, en prenant $\varepsilon_v = \varepsilon$ et $K_v = K \cdot \| v \|$, nous obtenons que F_v est localement Lipschitzienne en x_o.

□

Propriété A.4 : Soit ϕ une fonction continûment différentiable définie sur $\mathbb{R}^p$ et à valeurs réelles. Si la fonction F est localement Lipschitzienne en x_o, alors $\partial(\phi \circ F)(x_o) = [\mathcal{J}F(x_o)]^T \cdot \nabla\phi(F(x_o))$.

Pour démontrer cette propriété, nous avons besoin du lemme suivant.

Lemme A.4.1 : Soient F une fonction définie sur $\mathbb{R}^p$ et à valeurs dans $\mathbb{R}^q$ et g une fonction définie sur $\mathbb{R}^q$ et à valeurs réelles.

Supposons que F soit localement Lipschitzienne au point x_o de $\mathbb{R}^p$ et que g soit localement Lipschitzienne en $F(x_o)$.

Notons $(g \circ F)$ par f.

Alors *i)* f est localement Lipschitzienne en x_o,

ii) $\partial f(x_o)$ est inclus dans

$$co \{ [\mathfrak{J}F(x_o)]^T \cdot \partial g(F(x_o)) \}.$$

De plus, grâce à cette propriété A.4, nous pouvons considérer le cas particulier qui suit.

Cas particulier A.4.2 :

Si F est une fonction localement Lipschitzienne en x_o, alors $\partial F_v(x_o) = [\mathfrak{J}F(x_o)]^T \cdot v$.

En effet, il suffit d'appliquer la propriété A.4 avec $\phi = \langle \cdot, v \rangle$ car cette application est continûment différentiable.

Preuve de la propriete :

Dans, la démonstration, nous utliserons le résulat suivant.

(R) Supposons que F soit localement Lipschitzienne sur $\mathcal{O}$, et soient x et y deux poinys de $\mathcal{O}$,

alors $F(y) - F(x)$ appartient à $co\, \mathfrak{J}F([x,y]) \cdot (y - x)$.

La démonstration de ce résultat se trouve dans la référence [2] à la page 72.

1°) Montrons que $\partial(\phi \circ F)(x_o)$ est inclus dans $[\Im F(x_o)]^T \cdot \nabla\phi(F(x_o))$.

Comme ϕ est une fonction continûment différentiable , elle est localement Lipschitzienne. Nous pouvons donc appliquer le lemme A.4.1 et nous obtenons que

$\partial(\phi \circ F)(x_o)$ est contenu dans co $\{ [\Im F(x_o)]^T \cdot \partial\phi(F(x_o)) \}$.

Par la propriété 2.2.3, nous savons que $\partial\Phi(F(x_o)) = \{\nabla\phi(F(x_o))\}$.

Ainsi, $\partial(\phi \circ F)(x_o)$ est inclus dans co $\{ [\Im F(x_o)]^T \cdot \nabla\phi(F(x_o)) \}$.

Or, $[\Im F(x_o)]^T \cdot \nabla\phi(F(x_o))$ est un ensemble convexe.

En effet, soient x et y appartenant à $[\Im F(x_o)]^T \cdot \nabla\phi(F(x_o))$.

Par définition de $[\Im F(x_o)]^T \cdot \nabla\phi(F(x_o))$, nous savons qu'il existe des matrices M et N de $\Im F(x_o)$ telles que $x = M^T \cdot \nabla\phi(F(x_o))$ et $y = N^T \cdot \nabla\phi(F(x_o))$.

Soit λ un nombre réel compris entre 0 et 1. Nous calculons que $(\lambda \cdot x + (1-\lambda) \cdot y) = (\lambda \cdot M + (1-\lambda) \cdot N)^T \cdot \nabla\phi(F(x_o))$. La définition 3.1.1 assure que $\Im F(x_o)$ est convexe donc $(\lambda \cdot M + (1-\lambda) \cdot N)$ appartient à $\Im F(x_o)$.

D'où $(\lambda \cdot x + (1-\lambda) \cdot y)$ appartient à $[\Im F(x_o)]^T \cdot \nabla\phi(F(x_o))$, c'est-à-dire que $[\Im F(x_o)]^T \cdot \nabla\phi(F(x_o))$ est un ensemble convexe.

D'où co $\{ [\Im F(x_o)]^T \cdot \nabla\phi(F(x_o)) \} = [\Im F(x_o)]^T \cdot \nabla\phi(F(x_o))$.

Donc $\partial(\phi \circ F)(x_o) \subseteq [\Im F(x_o)]^T \cdot \nabla\phi(F(x_o))$.

2°) Montrons que $[\Im F(x_o)]^T \cdot \nabla\phi(F(x_o))$ est inclus dans $\partial(\phi \circ F)(x_o)$.

Il suffit de montrer que $[\Im(F, x_o)]^T \cdot \nabla\phi(F(x_o))$ est inclus dans $\partial(\phi \circ F)(x_o)$ car $\partial(\phi \circ F)(x_o)$ est un ensemble convexe.

Soit x appartenant à $[\mathfrak{J}(F,x_o)]^T\cdot\nabla\phi(F(x_o))$. Ainsi, il existe une matrice M appartenant à $\mathfrak{J}(F,x_o)$ telle que $x = M^T\cdot\nabla\phi(F(x_o))$. Par définition de $\mathfrak{J}(F,x_o)$, il existe une suite $(x_n)_{n=1}^{\infty}$ de points de Ω_F^c qui converge vers x_o telle que $M = \lim_{n\to\infty} JF(x_n)$.

D'où, $x = \lim_{n\to\infty} [JF(x_n)]^T\cdot\nabla\phi(F(x_o))$.

Or, le composé de fonctions différentiables est différentiable, il suit que, $(\phi \circ F)$ est différentiable en x_n et les règles de différentiation entraînent que

$$\nabla(\phi \circ F)(x_n) = [JF(x_n)]^T\cdot\nabla\phi(F(x_o)).$$

D'où, il existe une suite $(x_n)_{n=1}^{\infty}$ de points de $\Omega_{\phi\circ F}^c$ qui converge vers x_o telle que $x = \lim_{n\to\infty} \nabla(\phi \circ F)(x_n)$, c'est-à-dire que x appartient à $\partial(\phi \circ F)(x_o)$.

□

Preuve du lemme :

i) Montrons que f est localement Lipschitzienne en x_o.

Nous devons obtenir qu'il existe des constantes $\varepsilon > 0$ et $K > 0$ telles que $|f(x) - f(y)| \le K\cdot\|x - y\|$ pour tout x et y appartenant à $B(x_o,\varepsilon)$.

Comme g est localement Lipschitzienne en $F(x_o)$, il existe des constantes $\varepsilon_g > 0$ et $K_g > 0$ telles que

$|g(x_1) - g(y_1)| \le K_g\cdot\|x_1 - y_1\|$ quels que soient x_1 et y_1 dans $B(F(x_o),\varepsilon_g)$.

Et F étant continu en x_o (car F est localement Lipschitzienne en x_o), il existe $\varepsilon^* > 0$ tel que x appartient à $B(x_o,\varepsilon^*)$ entraîne que F(x) appartient à $B(F(x_o),\varepsilon_g)$.

Et le caractère localement Lipschitzien de F en x_o implique qu'il existe des constantes $\varepsilon_F > 0$ et $K_F > 0$ telles que

$$\| F(x) - F(y) \| \leq K_F \cdot \| x - y \| \text{ pour tout } x \text{ et } y \text{ dans } B(x_o, \varepsilon_F).$$

Prenons $\varepsilon = \min(\varepsilon^*, \varepsilon_F)$.

Or, par définition de f, nous savons que

$$|f(x) - f(y)| = |g(F(x)) - g(F(y))|.$$

Ainsi, pour tout x et y appartenant à $B(x_o,\varepsilon)$, nous obtenons que

$$|f(x) - f(y)| = K_g \cdot K_F \cdot \| x - y \|.$$

Donc, en prenant $K = K_g \cdot K_F$, f est une fonction localement Lipschitzienne en x_o.

ii) Montrons que $\partial f(x_o)$ est inclus dans $co\{ [\mathfrak{J}F(x_o)]^T \cdot \partial g(F(x_o)) \}$.

Soit u dans $\mathbb{R}^p$. La propriété 2.2.4 indique que

$$f^\circ(x_o;u) = \max \{ \langle x,u \rangle \mid x \in \partial f(x_o) \}.$$

Supposons qu'il existe un point y de $\partial g(F x_o))$ et une matrice M de $\mathfrak{J}F(x_o)$ tels

$$f^\circ(x_o;u) \leq \langle y, Mu \rangle.$$

Ainsi,

$$\max \{ \langle x,u \rangle \mid x \in \partial f(x_o) \} \leq \langle y, Mu \rangle.$$

Comme y appartient à $\partial g(F(x_o))$ et que M appartient à $\mathfrak{J}F(x_o)$, il suit que

$$M^T \cdot y \text{ appartient à } co\{ [\mathfrak{J}F(x_o)]^T \cdot \partial g(F(x_o)) \}.$$

D'où,

$$\max \{ \langle x,u \rangle \mid x \in \partial f(x_o) \} \leq \sup \{ \langle z,u \rangle \mid z \in co\{ [\mathfrak{J}F(x_o)]^T \cdot \partial g(F(x_o)) \} \}. \tag{An. 1}$$

Or, par la propriété 2.2.2, $\partial f(x_o)$ et $\partial g(F(x_o))$ sont des ensembles convexes compacts. D'où, $\partial f(x_o)$ et $\partial g(F(x_o))$ sont des ensembles convexes fermés. De plus, grâce à la propriété 3.1.2, nous savons que $\mathfrak{J}F(x_o)$ est un ensemble convexe compact, donc $\mathfrak{J}F(x_o)$ est un ensemble convexe fermé.

Ainsi, co { $[\mathfrak{J}F(x_o)]^T \cdot \partial g(F(x_o))$ } est un ensenble convexe fermé car la propriété 1.7.7 nous indique que l'enveloppe convexe d'un ensemble compact est compacte et l'enveloppe convexe d'un ensemble est convexe par définition.

Finalement, comme co { $[\mathfrak{J}F(x_o)]^T \cdot \partial g(F(x_o))$ } et $\partial f(x_o)$ sont des ensembles convexes fermés et que l'inégalité (An. 1) est vérifiée, la propriété 1.7.4 entraîne que

$\partial f(x_o)$ est contenu dans co { $[\mathfrak{J}F(x_o)]^T \cdot \partial g(F(x_o))$ }.

En <u>conclusion</u>, il faut donc montrer qu'il existe un point y de $\partial g(F(x_o))$ et une matrice M de $\mathfrak{J}F(x_o)$ tels que

$$f^\circ(x_o;u) \leq \langle y, Mu \rangle.$$

Comme g est localement Lipschitzienne en $F(x_o)$, il existe des constantes $\varepsilon_g > 0$ et $K_g > 0$ telles que

$|g(x_1) - g(y_1)| \leq K_g \cdot \|x_1 - y_1\|$ quels que soient x_1 et y_1 dans $B(F(x_o),\varepsilon_g)$.

De plus, F étant continu en x_o (car F est localement Lipschitzienne en x_o), il existe $\varepsilon^* > 0$ tel que x apaprtient à $B(x_o,\varepsilon^*)$ entraîne que F(x) appartient à $B(F(x_o),\varepsilon_g)$.

Considérons un point x de $B(x_o,\varepsilon^*)$. Ainsi, F(x) et $F(x+\lambda u)$ appartiennent à $B(F(x_o),\varepsilon_g)$ pour λ assez petit.

Nous allons montrer qu'il existe v dans $[F(x),F(x+\lambda u)]$, ξ dans $\partial g(v)$ et M dans $\mathfrak{J}F(\ [x,x+\lambda u]\ )$ pour lesquels

$$\frac{f(x+\lambda u) - f(x)}{\lambda} \leq \langle \xi, Mu \rangle,$$

où $[F(x),F(x+\lambda u)] = \{\ z \in \mathbb{R}^q \mid \exists\ \eta$ un nombre réel compris entre 0 et 1 tel que $z = \eta\cdot F(x) + (1-\eta)\cdot F(x+\lambda u)\ \}$,

$[x,x+\lambda u] = \{\ z \in \mathbb{R}^p \mid \exists\ \eta$ un nombre réel compris entre 0 et 1 tel que $z = \eta\cdot x + (1-\eta)\cdot (x+\lambda u)\ \}$,

et $\mathfrak{J}F([x,x+\lambda u]) = \{\ M \in \mathcal{M}(q,p) \mid \exists\ t \in [x,x+\lambda u]$ tel que $M \in \mathfrak{J}F(t)\ \}\}$.

Ensuite, par un passage à la limite sur cette inégalité, nous obtiendrons y appartenant à $\partial g(F(x_o))$ et M appartenant à $\mathfrak{J}F(x_o)$ tels que

$$f^{\circ}(x_o;u) \leq \langle y, Mu \rangle.$$

1°) Montrons qu'il existe v dans $[F(x),F(x+\lambda u)]$, ξ dans $\partial g(v)$ et M dans $\mathfrak{J}F([x,x+\lambda u])$ pour lesquels $\dfrac{f(x+\lambda u) - f(x)}{\lambda} \leq \langle \xi, Mu \rangle$.

Appliquons le théorème 2.2.8 de la valeur intermédiaire à g sur l'intervalle $[F(x),F(x+\lambda u)]$. Ainsi, il existe v appartenant à $[F(x),F(x+\lambda u)]$ tel que

$$g(F(x+\lambda u)) - g(F(x)) \text{ appartient à } \langle\ \partial g(v)\ ,\ F(x+\lambda u) - F(x)\ \rangle.$$

Autrement dit, il existe v dans $[F(x),F(x+\lambda u)]$ et ξ dans $\partial g(v)$ tels que

$$g(F(x+\lambda u)) - g(F(x)) = \langle\ \xi\ ,\ F(x+\lambda u) - F(x)\ \rangle.$$

D'où, il existe v dans $[F(x),F(x+\lambda u)]$ et ξ dans $\partial g(v)$ tels que

$$\frac{g(F(x+\lambda u)) - g(F(x)}{\lambda} = \frac{\langle\ \xi\ ,\ F(x+\lambda u) - F(x)\ \rangle}{\lambda},$$

c'est-à-dire que

$$\frac{f(x+\lambda u) - f(x)}{\lambda} = \frac{\langle \xi , F(x+\lambda u) - F(x) \rangle}{\lambda}.$$

Or, grâce au résultat (R), nous savons que

$$\frac{F(x+\lambda u) - F(x)}{\lambda} \text{ appartient à co } \{ \mathfrak{J}F([x,x+\lambda u])\cdot u \}.$$

c'est-à-dire qu'il existe un point w de co { $\mathfrak{J}F([x,x+\lambda u])\cdot u$ }

tel que

$$\frac{F(x+\lambda u) - F(x)}{\lambda} = w.$$

De plus, la propriété 1.7.8 indique que

$$\sup \{ \langle \xi , z\rangle \mid z \in \text{co } \{ \mathfrak{J}F([x,x+\lambda u])\cdot u \} \}$$

$$= \sup \{ \langle \xi , z\rangle \mid z \in (\mathfrak{J}F([x,x+\lambda u])\cdot u) \}.$$

Ainsi, il existe une matrice M de $\mathfrak{J}F([x,x+\lambda u])$ tel que

$$\langle \xi , w\rangle \le \langle \xi , Mu\rangle,$$

c'est-à-dire que

$$\frac{f(x+\lambda u) - f(x)}{\lambda} \le \langle \xi , Mu\rangle.$$

Donc, nous avons trouvé v dans $[F(x),F(x+\lambda u)]$, ξ dans $\partial g(v)$ et M dans $\mathfrak{J}F([x,x+\lambda u])$ tels que

$$\frac{f(x+\lambda u) - f(x)}{\lambda} \le \langle \xi , Mu\rangle.$$

2°) Montrons qu'il existe y dans $\partial g(F(x_o))$ et M dans $\mathfrak{J}F(x_o)$ tels que

$$f^{\circ}(x_o;u) \le \langle y, Mu\rangle.$$

Considérons arbitrairement la suite $(x_n)_{n=1}^{\infty}$ de points de $\mathbb{R}^p$ qui converge vers x_o et la suite $(\lambda_n)_{n=1}^{\infty}$ de nombres réels qui converge vers 0 par valeurs positives.

Ainsi, par le critère de réduction aux suites, il suit que

$$\limsup_{n \to \infty} \frac{f(x_n+\lambda_n u) - f(x_n)}{\lambda_n} = f^{\circ}(x_o;u).$$

De plus, la convergence des suites implique qu'il existe un nombre N de $\mathbb{N}_o$ tel que x_n et $(x_n+\lambda_n u)$ appartiennent à $B(x_o,\varepsilon^*)$ pour tout $n \geq N$.

Ainsi, comme g est une fonction localement Lipschitzienne sur l'intervalle $[F(x_n),F(x_n+\lambda_n u)]$, le point précédent de la démonstration peut être appliqué.

Nous obtenons alors qu'il existe v_n dans $[F(x_n),F(x_n+\lambda_n u)]$, ξ_n dans $\partial g(v_n)$ et M_n dans $\mathfrak{J}F([x_n,x_n+\lambda_n u])$ tels que

$$\frac{f(x_n+\lambda_n u) - f(x_n)}{\lambda_n} \leq \langle \xi_n, M_n u\rangle. \tag{An. 2}$$

Grâce à la propriété 2.2.9, nous savons que l'application ∂g est localement bornée. D'où, il existe une constante $K > 0$ telle que $\|\xi_n\| \leq K$. La suite $(\xi_n)_{n=1}^{\infty}$ étant bornée, le théorème de Bolzano-Weierstrass indique qu'il existe une sous-suite $(\xi_{t_n})_{n=1}^{\infty}$ extraite de $(\xi_n)_{n=1}^{\infty}$ qui converge vers un élément y de $\mathbb{R}^q$.

De plus, remarquons que comme F est continue en x_o et que v_n appartient à $[F(x_n),F(x_n+\lambda_n u)]$, la suite $(v_n)_{n=1}^{\infty}$ converge vers $F(x_o)$.

Donc, nous avons une suite $(v_n)_{n=1}^{\infty}$ convergente vers $F(x_o)$ et une suite $(\xi_{t_n})_{n=1}^{\infty}$ convergente vers un élément y de $\mathbb{R}^q$ telle que ξ_{t_n} appartient à $\partial g(v_{t_n})$. Ainsi, comme la propriété 2.2.9 entraîne l'application ∂g est semi-continue supérieurement, nous obtenons que y appartient à $\partial g(F(x_o))$.

De la même façon, en utilisant des propriétés similaires pour le Jacobien généralisé, il existe une sous-suite $(M_{t_n})_{n=1}^{\infty}$ extraite de $(M_n)_{n=1}^{\infty}$ qui converge vers un élément M de $\mathcal{M}(q,p)$. L'intervalle $[x_n, x_n+\lambda_n u]$ se réduisant à x_o pour n convergeant vers l'infini, nous obtenons que M est une matrice de $\mathfrak{J}F(x_o)$.

Finalement, par passage à la limite sur l'égalité (An. 2), nous obtenons que

$$f^{\circ}(x_o;u) \leq \langle y, Mu \rangle,$$

où y appartient à $\partial g(F(x_o))$ et M appartient à $\mathfrak{J}F(x_o)$.

Ce qui termine la démonstration du lemme A.4.1.

□

B) Deuxième partie.

Considérons la fonction ϕ définie sur $\mathbb{R}^n$ et à valeurs réelles et soit x appartenant à $\mathbb{R}^n$. Nous avons les deux propriétés suivantes.

Propriété B.1 : Soit la suite $(x_i)_{i=1}^{\infty}$ de points de $\mathbb{R}^n$ qui converge vers x,

alors $\limsup_{i\to\infty} \phi(x_i) \leq \limsup_{y\to x} \phi(y)$.

Preuve :

Soit $\varepsilon > 0$.

Comme la suite $(x_i)_{i=1}^{\infty}$ converge vers x, il existe un nombre N de $\mathbb{N}_o$ tel que $\|x_i - x\| < \varepsilon$ pour tout i > N.

Ainsi, $\{ x_i \mid i > N \}$ est inclus dans $\{ y \in \mathbb{R}^n \mid \|y-x\| < \varepsilon \}$.

D'où, $\sup_{i \geq N} \phi(x_i) \leq \sup_{\|y-x\|<\varepsilon} \phi(y)$.

Or, $\inf_{N^* \in \mathbb{N}_0} \sup_{i \geq N^*} \phi(x_i) \leq \sup_{i \geq N} \phi(x_i)$.

Donc, nous obtenons que, pour tout $\varepsilon > 0$,

$$\inf_{N^* \in \mathbb{N}_0} \sup_{i \geq N^*} \phi(x_i) \leq \sup_{\|y-x\|<\varepsilon} \phi(y),$$

c'est-à-dire que

$$\inf_{N^* \in \mathbb{N}_0} \sup_{i \geq N^*} \phi(x_i) \leq \inf_{\varepsilon > 0} \sup_{\|y-x\|<\varepsilon} \phi(y).$$

Autrement dit, nous avons que

$$\limsup_{i \to \infty} \phi(x_i) \leq \limsup_{y \to x} \phi(y),$$

car $\limsup_{i \to \infty} \phi(x_i) = \inf_{N^* \in \mathbb{N}_0} \sup_{i \geq N^*} \phi(x_i)$, et

$$\limsup_{y \to x} \phi(y) = \inf_{\varepsilon > 0} \sup_{\|y-x\|<\varepsilon} \phi(y).$$

□

Propriété B.2 : Soit c un point de $\mathbb{R}^n$. Les deux assertions suivantes sont équivalentes.

i) $\limsup_{y \to x} \phi(y) \leq c$

ii) pour toute suite $(x_i)_{i=1}^{\infty}$ de points de $\mathbb{R}^n$ qui converge vers x,

$\limsup_{i \to \infty} \phi(x_i) \leq c$.

Preuve :

1°) Par contraposition, montrons que

si $\limsup_{y \to x} \phi(y) \leq c$,

alors pour toute suite $(x_i)_{i=1}^{\infty}$ de points de $\mathbb{R}^n$ qui converge vers x, nous avons que $\limsup_{i \to \infty} \phi(x_i) \leq c$.

Il faut prouver que $\limsup_{y \to x} \phi(y) > c$ entraîne qu'il existe une suite $(x_i)_{i=1}^{\infty}$ de points de $\mathbb{R}^n$ convergente vers x telle que $\limsup_{i \to \infty} \phi(x_i) > c$.

Comme $\limsup_{y \to x} \phi(y) > c$, les propriétés de la borne inférieure et de la borne supérieure indiquent que, pour tout $\varepsilon > 0$, il existe un élément y de $B(x,\varepsilon)$ tel que $\phi(y) > c$.

Pour tout i dans $\mathbb{N}_0$, prenons $\varepsilon = \frac{1}{i}$. Ainsi, à chaque i de $\mathbb{N}_0$, il existe un élément x_i dans $B(x,\frac{1}{i})$ tel que $\phi(x_i) > c$.

La suite $(x_i)_{i=1}^{\infty}$ de points de $\mathbb{R}^n$ que nous obtenons converge vers x, car la suite $(\frac{1}{i})_{i=1}^{\infty}$ converge vers 0 et $0 \leq \|x_i - x\| \leq \frac{1}{i}$ quel que soit i.

De plus, comme $\phi(x_i) > c$ pour tout i appartenant à $\mathbb{N}_0$, il suit que $\limsup_{i \to \infty} \phi(x_i) > c$ car $\limsup_{i \to \infty} \phi(x_i)$ est une valeur d'adhérence (la plus grande) de la suite $(\phi(x_i))_{i=1}^{\infty}$.

2°) La seconde implication suit directement en appliquant la propriété B.1

□

Nous utiliserons également les propriétés qui suivent.

Lemme B.3: Soient $(\alpha_n)_{n=1}^{\infty}$ et $(\beta_n)_{n=1}^{\infty}$ deux suites de nombres réels. Supposons que la suite $(\alpha_n)_{n=1}^{\infty}$ converge vers un nombre réel a,

alors $\limsup_{n\to\infty} (\alpha_n + \beta_n) = a + \limsup_{n\to\infty} \beta_n$.

Preuve:

Par la propriété 1.3.3, nous savons que

$$\limsup_{n\to\infty} (\alpha_n + \beta_n) \leq a + \limsup_{n\to\infty} \beta_n.$$

Ainsi, il suffit de montrer que

$$\limsup_{n\to\infty} (\alpha_n + \beta_n) \geq a + \limsup_{n\to\infty} \beta_n.$$

Posons $b = \limsup_{n\to\infty} \beta_n$.

La propriété 1.3.2 indiquant que b est la plus grande valeur d'adhérence de la suite $(\beta_n)_{n=1}^{\infty}$, il existe une sous-suite $(\beta_{q_n})_{n=1}^{\infty}$ extraite de $(\beta_n)_{n=1}^{\infty}$ qui converge vers b.

D'où,

$$\limsup_{n\to\infty} (\alpha_{q_n} + \beta_{q_n}) = \lim_{n\to\infty} \alpha_{q_n} + \limsup_{n\to\infty} \beta_{q_n},$$
$$= a + b.$$

Autrement dit, (a + b) est une valeur d'adhérence de la suite $(\alpha_n + \beta_n)_{n=1}^{\infty}$, mais ce n'est pas nécessairement la plus grande.

D'où,

$$(a + b) \leq \limsup_{n\to\infty} (\alpha_n + \beta_n).$$

Finalement, nous obtenons que

$$\lim_{n\to\infty} \alpha_{q_n} + \limsup_{n\to\infty} \beta_{q_n} \leq \limsup_{n\to\infty} (\alpha_n + \beta_n).$$

Donc,

$$a + \limsup_{n\to\infty} \beta_n \leq \limsup_{n\to\infty} (\alpha_n + \beta_n).$$

□

Propriété B.4: Soient ψ et φ deux fonctions définies sur $\mathbb{R}^p$ et à valeurs réelles et soit a un point de $\mathbb{R}^p$.

Supposons que $\lim_{x\to a} \varphi(x)$ existe et vaut $\varphi(a)$,

alors $\limsup_{x\to a} (\psi + \varphi)(x) = \varphi(a) + \limsup_{x\to a} \psi(x)$.

Preuve:

Comme la propriété 1.3.5 indique que

$$\limsup_{x\to a} (\psi + \varphi)(x) \leq \varphi(a) + \limsup_{x\to a} \psi(x),$$

il suffit de montrer que

$$\limsup_{x\to a} (\psi + \varphi)(x) \geq \varphi(a) + \limsup_{x\to a} \psi(x).$$

Considérons une suite $(x_n)_{n=1}^{\infty}$ de points de $\mathbb{R}^p$ qui converge vers a. Par la propriété annexe B.1, nous savons que

$$\limsup_{n\to\infty} (\psi + \varphi)(x_n) \leq \limsup_{x\to a} (\psi + \varphi)(x).$$

Nous savons également que $\lim_{n\to\infty} \varphi(x_n) = \lim_{x\to a} \varphi(x) = \varphi(a)$.

Ainsi, le lemme B.3 entraîne que

$$\limsup_{n\to\infty} (\psi + \varphi)(x_n) = \varphi(a) + \limsup_{n\to\infty} \psi(x_n).$$

D'où,

$$\varphi(a) + \limsup_{n\to\infty} \psi(x_n) \leq \limsup_{x\to a} (\psi + \varphi)(x),$$

c'est-à-dire que

$$\limsup_{n \to \infty} \psi(x_n) \leq \limsup_{x \to a} (\psi + \varphi)(x) - \varphi(a).$$

Comme nous raisonnons sur une suite quelconque, la propriété B.2 entraîne que

$$\limsup_{x \to a} \psi(x) \leq \limsup_{x \to a} (\psi + \varphi)(x) - \varphi(a).$$

Donc

$$\varphi(a) + \limsup_{x \to a} \psi(x) \leq \limsup_{x \to a} (\psi + \varphi)(x).$$

□

C) Troisième partie.

Propriété C.1 : Toutes les matrices U de $\mathcal{M}(q,p)$ peuvent s'écrire comme $\sum_{k=1}^{r} a_k \otimes b_k$ où $r = \min(p,q)$, a_k appartient à $\mathbb{R}^p$, b_k appartient à $\mathbb{R}^q$ pour tout k compris entre 1 et r.

Preuve :

La démonstration va être anvisagée pour r = q uniquement car la preuve si r = p est similaire. Pour r = q, nous allons considérer les lignes de la matrices U. Dans le cas r = p, il suffit de prendre les colonnes de la matrices U.

Soit U une matrice de $\mathcal{M}(q,p)$.

Posons $U = [u_1, \ldots, u_q]^T$ où u_i appartient à $\mathbb{R}^p$ pour tout i compris entre 1 et q, et pour chaque i, posons $u_i = (u_{li})_{l=1}^{p}$.

De plus, pour tout j compris entre 1 et q, considérons les vecteurs e_j de $\mathbb{R}^q$ définis par $e_j = (\delta_{mj})_{m=1}^{q}$, où δ_{mj} est défini comme suit:

$$\delta_{mj} = \begin{cases} 0 \text{ si } m \neq j \\ 1 \text{ si } m = j \end{cases}.$$

Ainsi, pour tout k compris entre 1 et r, nous calculons que

$$u_k \otimes e_k = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ u_{1k} & \ldots & & u_{pk} \\ 0 & 0 & \ldots & 0 \end{bmatrix},$$

C'est-à-dire que

$$u_k \otimes e_k = (U_{mn}) \text{ avec } 1 \leq m \leq q \text{ et } 1 \leq n \leq p,$$

où $U_{mn} = \begin{cases} 0 & \text{si } m \neq k \\ u_{mn} & \text{si } m = k \end{cases}.$

En procédant ainsi, nous parcourons toutes les lignes de la matrice U. Donc, en sommant, nous obtenons que

$$U = \sum_{k=1}^{r} u_k \otimes e_k \ .$$

Cette décomposition de la matrice U n'est pas nécéssairement unique. Mais comme nous avons construit un tel découpage, nous pouvons dire que pour $r = \min(p,q)$, il existe des vecteurs a_k de $\mathbb{R}^p$ et b_k de $\mathbb{R}^q$ tels que $U = \sum_{k=1}^{r} a_k \otimes b_k$.

□

D) QUATRIÈME PARTIE.

Nous rappelons ici le théorème 14 énoncé par R.T. ROCKAFELLAR dans [8] à la page 155 et qui concerne le problème dual utilisé à l'exemple 4.1.10.

Théorème D.1 : Dans le cas convexe avec $r > 0$,

alors *i)* g_r est continûment différentiable sur $\mathbb{R}^q$,

ii) ∇g_r est Lipschitzien de constante $(\frac{1}{2.r})$.

Autrement dit,

$$| \nabla g_r(y) - \nabla g_r(x) | \leq (\frac{1}{2.r}) \cdot | y - x |$$

pour tout x et y dans $\mathbb{R}^q$.

REFERENCES

[1] F.H. CLARKE

"On the inverse theorem".

Pacific Journal of Mathematics (volume 64, numéro 1, pages 97 à 102, 1976).

[2] F.H. CLARKE

"Optimization and nonsmooth analysis".

Wiley (chapitre 2, 1983).

[3] FLETCHER

"Practical methods of optimization".

J. Wiley & sons (chapitre 9, 1980).

[4] A. FRIEDMAN
"Differential games".
J. Wiley & Sons (chapitre 4, 1971).

[5] J.-B. HIRIART-URRUTY
"New concepts in nondifferential programming".
Journées d'Analyse Non Convexe (Université de Pau, 1977).

[6] J.-B. HIRIART-URRUTY
"Characterizations of the plenary hull of the generalized Jacobian matrix".
Math. Programming Study (volume 17, pages 1 à 12, 1982)

[7] J.-B. HIRIART-URRUTY, J.-J. STRODIOT et V.H. NGUYEN
"Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data".
Applied Mathematics and Optimization (volume 11, pages 43 à 56, 1984).

[8] R.T. ROCKAFELLAR
"Solving a nonlinear programming problem by way of a dual problem".
Symposia Mathematica (volume 19,pages 135 à 160, 1976).